



Corso Luigi Einaudi, 55 - Torino

Appunti universitari

Tesi di laurea

Cartoleria e cancelleria

Stampa file e fotocopie

Print on demand

Rilegature

NUMERO: 1979A -

ANNO: 2016

A P P U N T I

STUDENTE: Canonico

MATERIA: Elaborazione segni biomedici - Prof. Molinari

Il presente lavoro nasce dall'impegno dell'autore ed è distribuito in accordo con il Centro Appunti.

Tutti i diritti sono riservati. È vietata qualsiasi riproduzione, copia totale o parziale, dei contenuti inseriti nel presente volume, ivi inclusa la memorizzazione, rielaborazione, diffusione o distribuzione dei contenuti stessi mediante qualunque supporto magnetico o cartaceo, piattaforma tecnologica o rete telematica, senza previa autorizzazione scritta dell'autore.

**ATTENZIONE: QUESTI APPUNTI SONO FATTI DA STUDENTIE NON SONO STATI VISIONATI DAL DOCENTE.
IL NOME DEL PROFESSORE, SERVE SOLO PER IDENTIFICARE IL CORSO.**

TECNICHE DI DEMOISING

indipendentemente da come viene acquisito in spm biologico, nel mondo biologico il rumore (sia sotto alle strumentazioni e sia per il segnale di per se) \Rightarrow il spm da acquisire è affetto da un livello di rumore t.c. l'elaborat. di quel segnale è complicata \Rightarrow "ripulire" il segnale, ed effettuare DEMOISING è un'operazione fondamentale.

Per rumore si intende tutto ciò ^{che} non è segnale, che non mi interessa. e ad es. il segnale che mi interessa è il segnale ECG \Rightarrow quando tocco gli elettrodi per il corpo, questi elettrodi non misurano solo l'attività elettrica del cuore, ma misurano di tutto \Rightarrow sia il spm ECG che non \Rightarrow formo questo modello di rumore additivo, che mi chiamo:

- $x(t) \rightarrow$ il segnale prelevato (quello che effettivamente abbiamo e disponiamo);
- $s(t) \rightarrow$ il nostro segnale biologico di interesse
- $n(t) \rightarrow$ rumore additivo, così temporalmente si sovrappongono il segnale utile

Alloca: $x(t) = s(t) + n(t)$

Questo però non è l'unico modello di rumore usato (è il più usato, ma non l'unico) infatti esistono casi in cui il rumore additivo non modella bene quello che avviene nella realtà e allora si ricorre a un modello di rumore moltiplicativo, per cui alcune volte si ha: $x(t) = s(t) \cdot n(t) \rightarrow$ ma questo succede in casi particolari.

$n(t)$ concettualmente è tutto quello che registriamo/acquisisco che non è spm che mi interessa; in pratica dipende da cosa e cosa perché è ovvio che ogni spm biologico è prevalentemente corrotto da un determinato tipo di rumore.

Es. semplice: se con elettrodi acquisiamo il spm ECG, stiamo acquisendo di certo anche l'attività elettrica dei muscoli respiratori (perché sono lì vicino...) \Rightarrow nel caso di acquisizione di spm ECG, $s(t)$ è il spm vero e proprio, mentre $n(t)$ è tutta l'attività elettrica che gli elettrodi registrano, ma che non è dovuta all'attività cardiaca (muscoli respiratori, tremore muscolare, potenziali superficiali, ecc...).

In un altro caso, se facciamo ^{col 23}registrazione di potenziali evocati, molto degli elettrodi nella zona che mi interessa (visiva, uditiva, temporale, ecc...) registriamo il potenziale evocato, ma registriamo anche un sacco di altre attività elettrica, che è quella degli altri neuroni, che normalmente chiamo segnale ECG \Rightarrow in questo caso il spm ECG perso, non è il spm biologico, ma è rumore perché il mio spm di interesse è il potenziale evocato

Quindi direi cosa è il rumore dipende da che tipo di suono registriamo e da che tipo di registrazione stiamo facendo.

Se prendiamo il segnale ECG fra lo scapolo e il polso, non sento i muscoli respiratori, ma si sente l'attività elettrica dei muscoli cardiaci che non è mai a riposo e riposo \Rightarrow \int sempre un tono muscolare di fondo, che crea un'attività elettrica $\neq 0$, che chiamiamo THRESHOLD e che si sovrappone a quello che andiamo a registrare.

Poi ci possono essere altri disturbi, come ad es. l'INTERFERENZA DI RETE \rightarrow non si può ottenere un'attenuazione ∞ , ma si può avere un CMRR sufficientemente alto, ma non infinito \Rightarrow per qualunque cosa succede durante la registrazione, ci può essere un po' di interferenza di rete che va nel segnale, cioè va in $n(t)$.

Quindi non è facile dire subito cosa è $n(t)$ \rightarrow bisogna conoscere: il sistema fisiologico che andiamo ad indagare, il sistema di prelievo, ecc... Se non conosco niente, il modello classico che si fa è quello di dire: il mio rumore che esocome il segnale è un processo casuale gaussiano bianco. Questo è un modello generico (scritto in prima battuta...), ma non descrive molto bene la situazione reale; però può essere un punto di partenza.

Un processo casuale \rightarrow è un segnale che ha una forma d'onda temporale, per ^{me} non prevedibile e quindi ha delle oscillazioni casuali. Un processo casuale è gaussiano (cioè processo casuale con distribuzione delle ampiezze di tipo gaussiano) se prendendo tutti i campioni di $n(t)$ e facendo un istogramma, in funzione del valore di tutti i campioni, e vedo quanti ce ne sono di ogni valore, si ottiene una distribuzione che è tipo una gaussiana \Rightarrow **SEGNALE GAUSSIANO** (o processo casuale gaussiano) non vuol dire processo casuale che inizia piccolo, poi diventa grande poi ritorna piccolo e fa lo stesso nel tempo, ma vuol dire che se prendo tutti i campioni di cui è fatto e ne faccio un istogramma, questo istogramma è fatto \sim come una gaussiana.

Bianco \rightarrow si dice un processo casuale bianco quando è totalmente decorrelato: ciò vuol dire che i campioni del segnale (in questo caso del rumore), non sono correlati con nessun altro campione, tranne che con se stessi. Allora un campione è correlato solo con se stesso, ma con nessun altro appartenente allo stesso processo casuale.

La funzione di autocorrelazione è una funzione che misura la correlazione fra i campioni di un segnale \Rightarrow se ho un segnale in cui ogni campione è correlato solo con se stesso, la sua funzione di autocorrelazione è una δ (DELTA DI DIRAC, nell'origine).

- 2) sgn e rum. sono segnali deterministici, entrambi;
- 3) sgn deterministico, rumore casuale
- 4) sgn casuale, rumore deterministico

Queste sono le 4 combinazioni possibili.

Nei 4 casi si utilizzano definizioni del rapporto sgn-rum di tipo differente perché la natura del sgn è differente.

CASO 1): sia rumore che sgn sono processi casuali.

In questo caso il rapporto sgn-rum. si definisce nel modo più semplice possibile come due definizioni: rapporto di POTENZE, allora si calcolano la potenza del sgn, la pot. del rumore e si fa:

$$SNR = \frac{P_{\text{segnale}}}{P_{\text{rumore}}} \quad (\text{in termini lineari})$$

Per convenire in dB, essendo un rapporto di potenze, si fa:

$$SNR_{dB} = 10 \cdot \log_{10} \frac{P_s}{P_n}$$

Per un processo casuale però come si calcola la potenza? ~~non si calcola~~
 Per un processo casuale la potenza non è altro che proporzionale alla varianza \rightarrow così per un processo casuale la stima della potenza associata al sgn è la varianza del processo casuale stesso.

\Rightarrow In questo caso di sgn e rum. processi casuali, il rapporto sgn-rumore è definito come il rapporto tra le varianze \rightarrow dato che la varianza è la deviazione standard al quadrato, per le proprietà del log, si può scrivere:

$$SNR_{dB} = 10 \cdot 2 \cdot \log_{10} \frac{\sigma_s}{\sigma_n} = 20 \log_{10} \frac{\sigma_s}{\sigma_n}$$

~~non si calcola~~

N.B.
 La conversione in dB è: 10 volte il ~~rapporto~~ \log_{10} del rapporto di potenze e 20 volte il \log_{10} del rapporto di ampiezza.

POSSIBILE PROBLEMA: se ho una registrazione in cui sono sovrapposti nel tempo sgn e rum come calcolo separatamente la P_s e la P_n ? Se avessi segnale da solo non dovrei fare né filtri, né demasking, né calcolare SNR, ma il problema è che molto spesso sgn e rum. sono sovrapposti nel tempo \Rightarrow calcolo del SNR

Caso 3: s_{gn} deterministico (quasi periodico, perché caso classico è il s_{gn} ECG, che oltre ad essere det. è quasi periodico) e rumore è un processo casuale \rightarrow in nel caso di ECG costituito da ~~rumore~~ ~~muscolare~~.

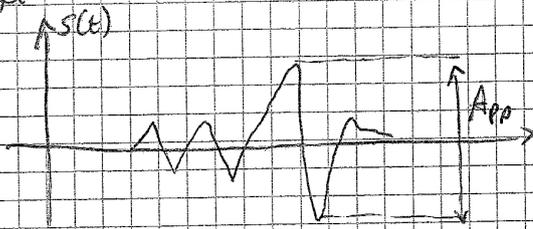
In questo caso quando il s_{gn} è det. (ha una forma d'onda precisa) e il rumore è casuale, conviene invece che un rapporto di potenze, fare un rapporto ampiezze \rightarrow così definire il SNR come rapporto di ampiezze \rightarrow è questo punto su già che quando lo vorrà convertire in dB, dovrà fare: $20 \log_{10}(\dots)$.

Quindi: $SNR = \frac{A_s}{A_n}$, con $\left\{ \begin{array}{l} A_s = \text{ampiezza } s_{gn} \\ A_n = \text{ampiezza rum} \end{array} \right.$

Ma come si misurano?

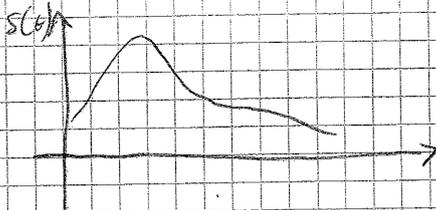
qualcosa

l'ampiezza di s_{gn} è facile: suppongo che il s_{gn} $s(t)$ sia un potenziale qual fatto così



La stima dell'ampiezza è l'ampiezza picco-picco (A_{pp}) (l'amp. max - l'amp. min)

Se avessi un potenziale più semplice, fatto così:



\Rightarrow Dato che è solo positivo, allora prendere solo l'ampiezza di picco

Quindi: prendere l'ampiezza di picco o di picco-picco, dipende da come è fatto il potenziale. Questo è lo misurava dell'ampiezza del segnale.

l'ampiezza di rumore è più faticosa ~~definita~~ ~~definita~~ \rightarrow calcolare la potenza è immediato: la potenza è la varianza.

Siccome l'ampiezza ha rispetto alle potenze un fattore $[20]$ o moltiplicatore, mentre la pot. la moltiplica per $[40]$ quando si fa il $\log \rightarrow$ se la varianza è una misura della potenza del rumore, la dev. standard è una misura della sua ampiezza \Rightarrow $SNR = \frac{A_{pp}}{\sigma_n}$ \rightarrow ma questa def. del punto di vista

statistico, non è corretto perché la varianza statisticamente è def. come somma degli scarti quadratici rispetto al valore medio \rightarrow tiene conto di tutta variabilità che c'è nei campioni. Ma la dev. standard invece dipende da lei è data la distribuzione di tutta var. casuale. Variabili casuali con la A_{pp}

Nel nostro caso nessuno non si sa se $\sigma_N = 1$, ~~ma~~ per cui si scrive: $4 \cdot \sigma_N$.

ESERCIO

questi det. corretto da un rumore che è un proc. casuale

Allora mi danno un segnale e mi chiedono di calcolare il SNR e lo definisco come: $SNR = \frac{A_s}{\sigma_N}$. È un errore? NO, ma l'importante è che sia

chiaro come è stato definito (in questo caso sto assumendo il rumore).

Se voglio l'ora al 98% (e non al 95%) \Rightarrow non sarà $4 \cdot \sigma$, ma sarà 6σ o 7σ , va bene \rightarrow è una convenzione e convenzionalmente noi accettiamo un errore di

1 specie del 5%; ma se uno è più "restrittivo", va bene \rightarrow si può mettere un errore di 1 specie piccolo e piccolo, modificando la definizione di SNR \Rightarrow l'importante è che sia chiaro come è stato definito. L'unica cosa che non deve succedere è quella di dare un valore di SNR, facendo rimanere la sua definizione misteriosa. In pratica va bene definirlo con il σ , ma se non si def. con il σ ma è un errore grave.

Allora il caso classico è: un segnale ECG corretto da un processo casuale che può essere tipo il tremore muscolare \rightarrow per ~~un~~ segnale ECG ha più senso A_{pp} come misura dell'ampiezza del segnale (perché c'è un'importante parte negativa).

Allora si scopre che: (SCI 12)

$$A_{pp} = 670 \text{ u.a.} \quad ; \quad \sigma_N = 57 \text{ u.a.} \quad \Rightarrow \quad SNR = \frac{670}{4 \cdot 57} = 294 \quad \text{oppure: } SNR_{dB} = 9,36 \text{ dB}$$

UNITÀ ARBITRARIA

\hookrightarrow perché l'ora non è totale (né in mV, né in nessuna unità...)

\Rightarrow dato che: $SNR < 10 \text{ dB} \Rightarrow$ il segnale non è un granché di SNR.

NOTA

Nella SCI 12 vediamo stare il segnale scuro che sarebbe il segnale puro, senza rumore (impossibile, ma facciamo finta che sia così...) \Rightarrow in questo caso il SNR sarebbe INFINITO.

Nella pratica abbiamo solo il segnale verde, per cui calcoliamo l'ampiezza picco-picco del segnale e la ~~dev. stand.~~ dev. standard del rumore e calcoliamo dove non c'è ECG \rightarrow cioè devo prendere un pezzo del tracciato verde dove non c'è il segnale ECG, perché altrimenti calcoliamo sia la dev. stand. del rumore che la dev. stand. del segnale \Rightarrow devo ricordare che il segnale ECG è fatto da un insieme di onde, poi c'è un tratto ~~isoelettrico~~ "isoelettrico" (dove non c'è niente), prima che compare il battito successivo \rightarrow in quel tratto, calcoliamo la dev. standard.

Come migliorare il SNR? È una regola che è quella che dà il nome a tutte le tecniche che vedremo: TECNICHE DI DENOISING → vuol dire che il segnale non si tocca! Se ho infatti un SNR insoddisfacente, si lavora solo sul denominatore, il segnale non si tocca, è quello che è. (Si chiama DENOISING, perché per ↑ SNR, bisogna ↓ (diminuire) il denominatore).

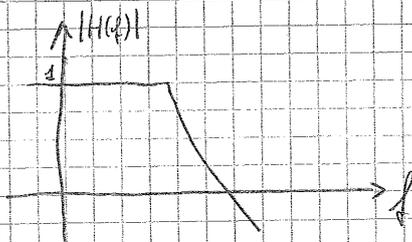
Noi li conosciamo come FILTRI, ma effettivamente si dicono metodi di denoising perché riducono la quantità di rumore che affligge il segnale.

Noi conosciamo i filtri: PASSA-ALTO, PASSA-BASSO, PASSA-BANDA, RIGETTA BANDA e PASSA-TUTTO → di questi, sappiamo il modulo (cioè il modulo della funzione di trasferimento di questi filtri) ma ci interessa sapere anche la fase → infatti si vedrà che ci sono filtri P-A, P-B, ecc. alcuni con fase lineare ed altri che non hanno fase lineare ⇒ questo sarà un problema e dovrà essere considerato.

!! Usiamo MATLAB che fa i diagrammi di Bode al posto nostro.

Allora: la teoria dei filtri digitali prevede che noi dobbiamo essere in grado di implementare il filtro, poi si fa tutto e Matlab e ci fa le curve e disegna le maschere (ma il filtro lo dobbiamo dimensionare noi, cioè noi dobbiamo occuparci del lavoro o monte di lavoro come è fatto il filtro, cioè dobbiamo capire perché dimensionare il filtro di un tipo o di un altro).

Tutti i filtri per definizione, in banda passante, hanno guadagno unitario:



Esempio di filtro P-BASSO

e noi vogliamo lavorare con filtri che in banda passante fanno passare, non amplificano, non fanno niente ⇒ guadagno unitario proprio perché le tecniche che usiamo devono lavorare sul rumore e non sul segnale → il segnale non si tocca.

I filtri P-ALTO, P-BASSO, P-BANDA e RIGETTA BANDA vanno bene quando si ha un rumore che affligge il segnale e che possiamo localizzare con facilità nell'area delle frequenze e trovare il filtro opportuno, toglierlo.

Quindi: CASO CLASSICO → nel dominio del tempo tutto è sovrapposto:

$$x(t) = s(t) + n(t)$$

Voglio migliorare il SNR, ottenendo il rumore medesimo; in frequenza vedo e vedere come è fatta la $\mathcal{F}\{x(t)\} = X(f)$ e scopro, ad esempio che:

sempre uguale.

- segnale EEG e interferenza di rete

Per disporre la trasformata di Fourier dell'interferenza di rete $N(f)$ → lo disegno come una δ a 50 Hz



Le bande di un μ EEG $\dot{\sim}$ da $0, \dots$ ^{questo} fino a 35-40 Hz \Rightarrow in questo caso fare DENOISING è banale: il filtro che ci serve è un filtro PASSA-BASSO. Così facendo il segnale non lo tocco proprio e $N(f)$ a 50 Hz se ne va.

OSS.

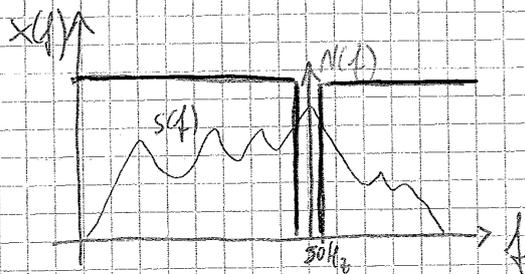
Non si usa un filtro riflette-banda perché questo è sempre più complicato del PASSA-ALTO e del PASSA-BASSO, perché ha 2 frequenze di taglio \Rightarrow la probabilità di errore raddoppia e quindi non conviene creare il riflette banda.

In generale dimensionare un filtro riflette-banda o passa-banda è sempre più difficile che dimensionare un passa-alto o un passa-basso.

A volte però, riflette-banda e passa-banda sono fondamentali, ad esempio quando l'interferenza di rete è esattamente in banda con il segnale. Allora, facciamo questo caso:

• Segn a banda larga

- Rumore è l'interferenza di rete a 50 Hz \rightarrow in banda stretta



Dato che noi il μ lo vogliamo tenere il minimo possibile \Rightarrow nel caso in cui μ e rumore sono sovrapposti in banda, ma ad esempio il μ è a banda larga e il rumore è a banda stretta (come in questo caso), allora

Quindi si adattano bene ad alcuni sgu, ma molto male ad altri.

⇒ I casi in cui non si può fare niente (sgu e rumore sono sovrapposti; sono due processi casuali, occupano più o meno le stesse bande, ...).

Pero, nel caso in cui ci siano sgu e rumore ~~che~~ che soddisfano determinate hp ⇒ qualcosa si può fare

La TECNICA DELL'AVERAGING è ^{quasi} il argomento della 1° esercitazione.

INTRODUZIONE ALLA **TECNICA DELL'AVERAGING**

La tecnica dell'averaging serve sempre e solo in presenza di sgu deterministici. Come detto, devono essere soddisfatte alcune hp:

1° hp: il sgu di interesse deve essere deterministico prevedibile.

In modo prevedibile, non vuol dire rispetto alle forme del sgu che deve essere prevedibile (quello è già compreso nel fatto che il sgu è det.), ma che io conosco per certo quando questo segnale c'è e quando non c'è.

ESEMPIO: il sgu ECG è un sgu deterministico, ma NON PREVEDIBILE (non posso sapere con certezza quando c'è un battito e quando non c'è)

⇒ i sgu classici che rispondono a queste hp di prevedibilità sono quelli detti POTENZIALI EVOCATI → sono detti così perché sono evocati, occorrono quando sono chiamati, cioè: è la risposta di un sistema fisiologico ad una stimolazione fornita dall'esterno. Se non c'è nessuna stimolazione dell'esterno, sono sicuro che quel potenziale non esiste, così come se ho stimolato, sono sicuro che il potenziale ci deve essere.

ES: Potenziale evocato di tipo uditivo: metto 2 elettrodi nella zona della superficie corticale (che è responsabile della funzione uditive ⇒ vicino all'orecchio). Quando stimolo il soggetto con uno stimolo acustico, deve risultare un pot. evoc. di tipo uditivo; quando non lo stimolo, non c'è.

Questo quindi vuol dire che si ripete in modo prevedibile.

⇒ i sgu che soddisfano queste hp si riducono ⇒ non è una tecnica sempre utilizzabile.

2° hp: il sgu deterministico si deve ripetere un n° suff. alto di volte, sostanzialmente identico a se stesso



ogni registrazione

Ogni registrazione (epoca) i , può essere descritta come:

$$x_i(t) = s(t) + m_i(t)$$

oss: $s(t) \rightarrow$ spm non dipende dall'epoca (per le hp)
 $m_i(t) \rightarrow$ $i \neq j$ epoca

Se calcolo il $\left(\frac{S}{N}\right)$ di quest'epoca i -esima (Problema: un solo rapporto di ampiezza o potenza per N ?
 Rapporto di segnale perché spm deterministico e rumore casuale)

$$SNR_{x_i} = \frac{A_s^2}{4 \sigma_N^2}$$

(\rightarrow Ampiezza pico-peak, ad es.)
 dev. std di calcolo dopo noi, cioè il dev. cui alla fine della registrazione.

$i = 1, \dots, 100$ (epoche)

oss: Se misuro il SNR di tutte le epoche e 100 le epoche, provo lo stesso valore (ovviamente): perché il spm è uguale a se stesso $\Rightarrow A_s = A_s$ (epoca); il rumore è 1/100 in senso lato, cioè vuol dire che la sua proprietà energetiche non variano nel tempo \Rightarrow ha lo stesso $\sigma^2 \Rightarrow$ ha lo stesso σ

Di queste $N (= 100)$ epoche, facciamo la:

$$\Rightarrow \bar{x}(t) = \frac{1}{N} \sum_{i=1}^N x_i(t) = s(t) + \frac{1}{N} \sum_{i=1}^N m_i(t)$$

Questo è lo medio di una variabile casuale

Quindi, ora calcolo SNR di $\bar{x}(t)$ e confrontarlo con SNR dell'epoca i -esima, x_i :

$$SNR_{\bar{x}} = \frac{A_s^2}{4 \frac{\sigma_N^2}{N}} = SNR_{x_i} \cdot \sqrt{N}$$

* SNR in linear, non in dB

oss: L'operazione di media ha causato il fatto che partivamo da una situazione iniziale in cui l'ampiezza del rumore era pari a σ_N e si è passati ad $\frac{\sigma_N}{\sqrt{N}}$ condizioni in cui l'ampiezza di rum $i = \frac{\sigma_N}{\sqrt{N}}$
 \rightarrow SNR moltiplica di un fattore \sqrt{N}

DOVE È STATO

Calcolata la varianza come: $E[m_i(t)]^2$, cioè:
 $E[\bar{x}(t) - s(t)]^2 = \frac{1}{N^2} E\left[\sum_{i=1}^N m_i(t)\right]^2 = \frac{1}{N^2} \sum_{i=1}^N E[m_i(t)]^2 = \frac{1}{N^2} \cdot N \cdot \sigma_N^2 = \frac{\sigma_N^2}{N}$

CONSIDERAZIONI:

cioè voglio ridurre il rumore

Se voglio SNR alto \Rightarrow devo aumentare $N \Rightarrow \sqrt{N}$, \sqrt{N} è l'aumento di SNR. Ho detto che N è limitato a 100/200, nel caso classico (biologico). È difficile avere $N=500/600$, va bene $\sim 130/140$

!! CLASSICA ORMANO D'ESAME: Dato un spm, calcolata SNR e poi protetto ed essere almeno 12 dB. So esattamente quante medie devo fare. Se ad es. ho fatto la registrazione di un potenziale evocato e mi accorgo che SNR è solo di 7 dB e lo voglio portare a 12/15 dB \Rightarrow io o prima grandi stimoli devo fornire.
 La cosa bella della tecnica dell'overlapping è che se le hp fatte all'inizio, sono rispettate, io riesco ad applicare la tecnica dell'overlapping avendo delle prestazioni note e precise, infatti si può di quanto moltiplica il SNR.
 Ovviamente vale anche il contrario, cioè quelle hp fatte sono condizioni necessarie e suff. e se una di quelle non è verificata, non è che la tecnica dell'overlapping non funziona più; forse un moltiplicamento del SNR o l'ho ancora, ma non so dire di quanto \Rightarrow perdo la capacità di sapere a priori di quanto moltiplicare il mio SNR.
 Un caso semplice è quando il rumore non è più stazionario in senso lato ma si ha più lo stesso dev. standard, quindi non si può sapere quale sarà lo dev. staz. finale del rumore; magari sarà anche minore rispetto a quella originale (e quindi ci guadagno sul SNR), però non so di quanto.

L'implementazione delle tecniche dell'averaging è offerta dallo 1° ESERCIZIO.
Avremo ugu che rispettano tutte hp alla 10 esecuzione → c'è un file
fido abbastanza completo in cui bisogna mettere le istruzioni che
mangiano per completare la tecnica dell'averaging.

L'unica cosa che bisogna fare è implementare la tecnica dell'averaging
sequenziale → vuol dire che bisogna controllare che non non che si
mediano delle epoche, il SNR cresce con una legge matematica nota
(quello esatto).

% commenti su Matlab

trovare più elementare del codice: dico "prendo le prime 2 epoche, le medio, al
ciclo successivo si sommo la terza epoca e divido per 3!" NON VA
BENE, perché le prime 2 le divido per 6 poi --

→ Quando si fanno le medie successive, prendo le prime 2 epoche, le
sommo e poi divido per 2; calcolo il SNR, prima di uscire del
ciclo moltiplico per 2, così ho solo la prima (non la media!)
al ciclo successivo sommo la 3^a, divido per 3, calcolo il SNR, moltiplico
per 3, esco dal ciclo; rientro nel ciclo, sommo la 4^a, divido per 4, calcolo
il SNR, moltiplico per 4 e così via --

OSS.

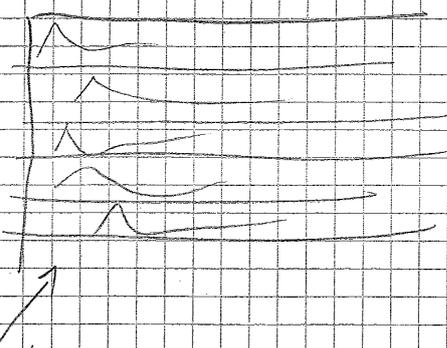
Se applico AVERAGING in assenza di jitter, il SNR migliora di un fattore
N.

zone media: $\text{mean}(A)$. NOTA Matlab lavora per colonne \Rightarrow se A è una matrice (130×100) , allora scrivendo $\text{mean}(A)$, mi viene restituito un vettore che ha tanti elementi quanti sono le colonne della matrice



ogni elemento contiene la media della colonna i -esima della matrice

\Rightarrow Quando si ha un'istruzione che funziona su matrici, Matlab la applica sempre per colonne. Se sono invece nel caso del JITTER, prendo gli stessi dati e li rimetto per righe (qui era una riga della matrice $A \Rightarrow$ succede che ho tutti i potenziali disallineati nelle righe, perché è jitter



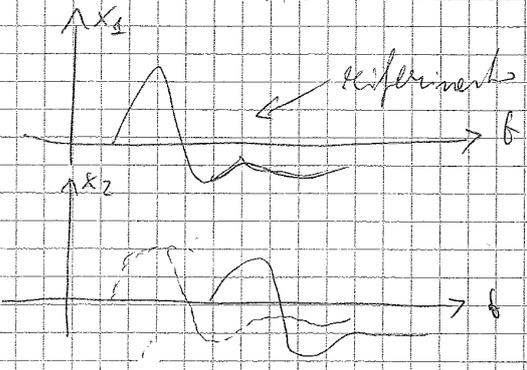
Allora: i potenziali non iniziano più tutti allo stesso campione e finiscono allo stesso campione, ma sono tutti disallineati.

i potenziali sono disallineati

Allora se faccio la media per colonne, non si fa più la media di cose uguali, ma essendo spostato il suo temporaneamente, anche lui ha una certa variabilità, che si può con un filtro P-BASSO.

Soluzione: Se il jitter ha spostato i potenziali, io li riallineo ma concretamente per fare ciò si prende un'epoca di riferimento e le altre $n-1$ epoch le riallineo rispetto quella di riferimento.

Cosa vuol dire riallineare? Facciamo un caso semplice: Ho 2 segnali uguali, solo che sono disallineati. Voglio riallineare x_2 su x_1 . Cosa devo fare? A uno dei (in questo caso x_2) devo applicare un ritardo in modo da portarlo ad essere, di nuovo, coincidente con x_1 . Quindi, per 1° caso devo calcolare il ritardo esistente tra x_1 e x_2 ; 2° caso devo applicare il ritardo calcolato al segnale per portarlo al punto voluto

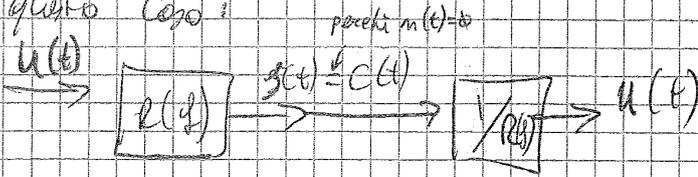


Lo è difficile a calcolare il ritardo tra 2 spn (2 epoch) che hanno un SNR = -4/-5. Questo è il problema.

\Rightarrow Si fa con un'istruzione funzione `delay`, che implementa una tecnica di riallineamento tra segnali che vedremo più avanti nel corso.

Se $m(t) = \delta$, la soluzione sarebbe semplice: prendo $c(t)$, faccio la deconvoluzione per R e ho la stima ottima $\Rightarrow U(f) = C(f)/R(f)$ [questo è facile nel dominio delle frequenze].

Ma in questo caso:



Il problema quindi non è $R(f)$, ma è il rumore che si inietta all'uscita del filtro. A noi il caso con $m(t) = \delta$ non interessa, ci interessa quando c'è il rumore. Quindi ci interessa dimensionare questo filtro, in presenza di rumore.



il segnale che c'è qua: $C(f) \cdot \Phi(f)$

$$\Rightarrow U'(f) = \frac{C(f) \Phi(f)}{R(f)}$$

Per capire cosa si intende per U' massimamente simile ad U oppure capire cosa vuol dire U ottimo rispetto ad U' , dobbiamo dare un criterio matematico: minimizzare l'errore quadratico \Rightarrow la stima U' sarà ottima, quando confrontata con U , produrrà un valore quadratico minimo.

DEF. Errore quadratico: prendo U , sottraggo U' dal quadrato e integro su tutta l'area. Se U fosse una stima perfetta, l'errore quadratico sarebbe $= 0$. Questo è il metodo corretto per definire cosa intendiamo come ottimo e non abbiamo definito (correttamente) il ottimo nel dominio del tempo, perché a noi interessa che U' sia uguale ad U nel ^{tempo} che produce un errore _{minimo}.

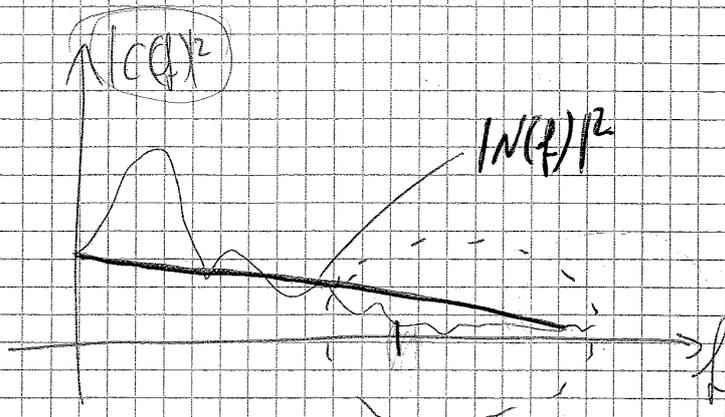
l'ottimo succede che si voglia ^{trovare} quel valore di $\phi(t)$ che produca un $U'(t)$ minimo \Rightarrow devo derivare l'errore rispetto a ϕ e porre $= 0$ e vedere se ϕ è un minimo. Nel tempo è complicato fare ciò (perché ϕ si esprime in nel tempo, ci sono conclusioni, cioè ipotesi in tempo \Rightarrow diventa complicato), per cui invece di minimizzare quella funzione nel tempo, lo faccio nelle freq. \Rightarrow trovare il minimo errore (nelle freq) e non nel tempo. Questo mi lo permette di fare Poessel \Rightarrow Uo Poessel e altro slide 27

quindi minimizzo:
$$e = \int_{-\infty}^{\infty} [U'(f) - U(f)]^2 df =$$

[ma che questo qua:]
$$e = \int_{-\infty}^{\infty} [u'(t) - u(t)]^2 dt$$

Ma se non ho un segnale intermittente, bisogna usare il

METODO GRAFICO



Compensando, in un grafico, il |C(f)|^2 (lo spettro di C) con il |N(f)|^2 (lo spettro di N) si ottiene un segnale intermittente. In questo grafico, infatti, ho: $|C(f)|^2 = |N(f)|^2 + |S(f)|^2$, e questo è lo spettro.

Si sa e precisa che:

$$|C(f)|^2 = |S(f)|^2 + |N(f)|^2$$

Nell'hp che S e N siano correlati, \Rightarrow Considero le parte di spettro di C(f) ad alte freq. \Rightarrow Cosa succede? Abbiamo visto che si

Non biologici sono limitati in banda \Rightarrow non \exists componenti ad alta f (al più kHz) in più se ho C(f) \Rightarrow è stato campionato \Rightarrow rispetto il th di Nyquist ($f_c \geq 2B$, appare che la freq. di camp. deve essere almeno 2 volte la freq. massima presente nel sp.).

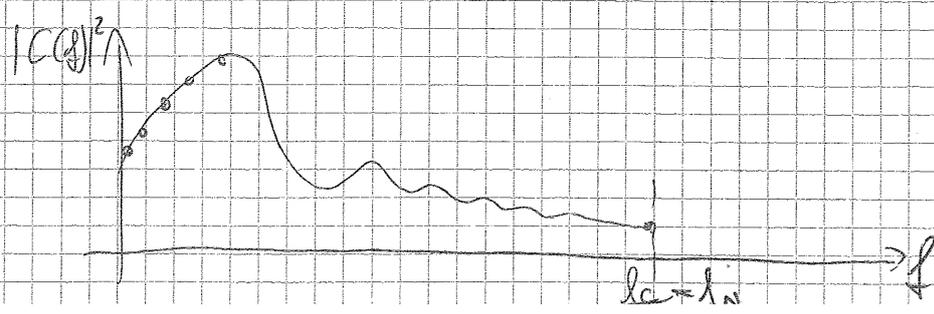
Se ad es. ho campionato un sp. EEG, che ha una banda fino a 35-40 Hz, fino a 30-40 Hz mi aspetto di trovare sp. più rumore, oltre 40 Hz trovo solo rumore, perché sp. è finito. Allora si può fare un metodo di interpolazione, e dire: considero lo spettro di C in alta frequenza e faccio finto che la code che vedo in alta freq. è solo rumore. Allora interpoliamo le caratteristiche (linee rosse nel grafico) dalle basse alle alte freq. e faccio finto che sia $= |N(f)|^2$. Ecco si ha:

oss: $|C(f)|^2 - |N(f)|^2 = |S(f)|^2$ \leftarrow da mettere al numeratore di $\Phi(f)$

Metodo grafico: grafico e approssimato (interpolaz. numer)

\Rightarrow lo uso solo nel caso di sp. non intermittenti, cioè se il primo metodo non funziona. Vediamo un esempio per essere:

Supponiamo C(f) un segnale campionato ad una f_c .

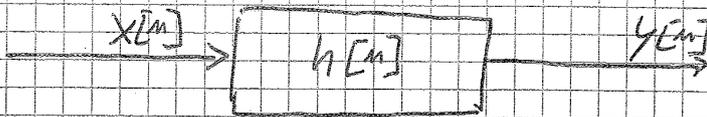


21/12

TRASFORMATA Z

Per esprimere come è la trasform. z, prendiamo come riferimento la trasform. di Fourier, che concettualmente è la descrizione del segnale in termini delle sue componenti armoniche, di una somma di sinusoidi complesse. Quindi prendo il segnale e non lo modifico, ma mi dà una descrizione della sua descrizione come serie di punti in funzione del tempo, ne dà una descrizione nella base di una serie di coeff. cerco del quale mi dice quanto di una determinata frequenza c'è all'interno del segnale.

La trasformata z è lo stesso caso, solo che invece di prendere come riferimento delle sinusoidi complesse, io scompongo il segnale secondo una base che è genericamente dipendente di un numero complesso. Supponiamo di avere un sistema LTI, una sequenza discreta $x[n]$ all'ingresso di questo sistema e all'uscita una sequenza $y[n]$ e supponiamo di chiamare $h[n]$ la risposta all'impulso di questo sistema



Se descriviamo cosa succede nel dom. del tempo di questo sistema, si può scrivere

$y[n] = x[n] * h[n]$ (La convoluzione nel discreto, invece dell'integrale, si fa sullo $\sum_{-\infty}^{+\infty}$)

$\Rightarrow y[n] = \sum_{k=-\infty}^{+\infty} x[n-k] h[k]$

Per definire la trasform. z, all'ingresso del filtro metto una $x[n]$ specifica cioè: $x[n] = z^n$, dove z è un n° complesso. Allora questa espressione

$y[n] = \sum_{k=-\infty}^{+\infty} x[n-k] h[k]$ posso riscriverlo: $y[n] = \sum_{k=-\infty}^{+\infty} z^{n-k} h[k]$

ma dato che: $z^{n-k} = z^n \cdot z^{-k}$, allora: $y[n] = z^n \sum_{k=-\infty}^{+\infty} z^{-k} h[k]$. Siccome z^n è esattamente la sequenza $x[n]$ d'ingresso, allora questa espressione $\sum_{k=-\infty}^{+\infty} z^{-k} h[k]$ è proprio la trasform. di $h[n]$. Cioè è la

✓ RISPOSTA IN FREQUENZA: $H(z) = \sum_{k=-\infty}^{+\infty} z^{-k} h[k]$ TRASF. Z DI UNA SEQUENZA $h[n]$

\Rightarrow La trasform. di Fourier è un caso particolare della trasformata z.

La trasform. z è una trasform. generica che scompone il segnale su una base complessa (di funzioni complesse), di cui la trasform. di Fourier è un sottocaso. Il sottocaso è quando

$z = e^{-j2\pi f}$

NB \Rightarrow quindi la t.d.f. \Rightarrow La trasform. z di una seq. discreta, è una quantità continua $\rightarrow []$ indicano una quantità discreta, $()$ indicano una quantità continua. PROPRIETÀ IMPORTANTI: la trasform. di una sequenza numerica discreta, è sempre una quantità continua in z, esotone come la f.d. di un n° discreto è una quantità continua in ω .

Calcolare la f. di t. e la risposta all'impulso complessiva che c'è tra y_2 e x_1 tramite di trasformata z.

SUBSTITUZIONE: per calcolare la f. di t., come già visto, si usa la sequenza $x[n] = z^n$ mentre per calcolare la risposta all'impulso equivalente, conviene usare come $x[n] = \delta$ (cioè un singolo impulso).

CALCOLO RISPOSTA ALL'IMPULSO: ($x[n] = \delta[n]$)

$$y_2[n] = \delta[n] * h_1[n] = h_1[n] ; y_2[n] = h_1[n] * h_2[n] \Rightarrow h_{eq}[n] = h_1[n] * h_2[n]$$

CALCOLO FUNZ. DI TRASFERIMENTO: ($x[n] = z^n$)

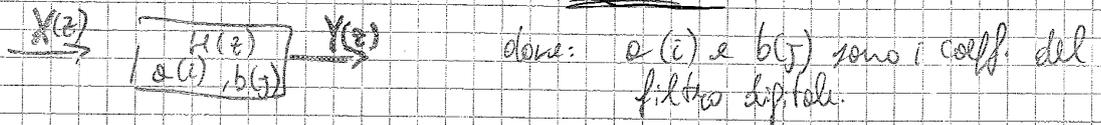
$$H_{eq}(z) = H_1(z) \cdot H_2(z) \quad \text{perché: } y_2[n] = x[n] * h_1[n] = z^n \sum z^{-k} h_1[k]$$

$$y_2[n] = y_1[n] * h_2[n] = z^n \sum z^{-k} h_1[k] \sum z^{-l} h_2[l]$$

$$\Rightarrow H_{eq}(z) = H_1(z) \cdot H_2(z)$$

ISLADES
 È più facile scapornare di filtri digitali in termini di transf. z che non in termini di f., ricordando che il passaggio tra $H(f)$ e $H(z)$ è semplicemente: $z = e^{-j2\pi f}$. Una volta definite le funz. di transf. nel piano z, il passaggio nel dominio di Fourier è immediato (per sostituzione di variabile). Ora è facile definire un **FILTRO DIGITALE**

⇒ Come è fatto un filtro digitale: è un st. LTI e supponiamo di avere:



È possibile scrivere genericamente:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{b(0) + b(1)z^{-1} + \dots + b(m_b+1)z^{-m_b}}{1 + a(1)z^{-1} + \dots + a(m_a+1)z^{-m_a}}$$

Qualunque filtro digitale si può scrivere così, in termini di rapporto di polinomi in funzione di z.
 • Il numeratore, è formato da una serie di coeff. b , che vanno da $b(0)$ o $b(m_b+1)z^{-m_b}$; cioè è un polinomio dove ci sono coeff. b che moltiplicano una serie di potenze discrete di z^{-1} (cioè $k=0, \dots, m_b$).
 • Al denominatore c'è la stessa cosa (invece che b , i coeff. a) e i poteri vanno da 0 a m_a ($k=0, \dots, m_a$).

Si può dimostrare che: mettendo i numeri opportuni all'interno dei coeff. b e selezionando i valori di m_b ed m_a opportuni, si possono avere fatte le funz. di transf. che normalmente si usano (P-ALTO, P-BASSO, P-BANDA, P-RETTA-BANDA, P-TUTTO) ⇒ qualunque filtro digitale può essere scritto come rapporto di 2 polinomi. Perché è possibile fare ciò? Per il motivo

se vediamo il numeratore, è un polinomio in funzione di z di ordine m_b quindi ammette m_b radici; dato che quelle radici annullano il numeratore della f. di t., sono gli zeri ⇒ il numero definisce il numero e la posizione degli zeri (il numero ottiene l'ordine m_b ; la posizione

Se ho: $H(z) = \frac{1 - z^{-5}}{1 - 3z^{-3}}$ Scrivere l'equazione è banale: il numeratore moltiplica li il peso, il denominatore moltiplica l'uscita

$\Rightarrow Y[n] = X[n] - X[n-5] + 3Y[n-3]$ # senza fare tutti i conti di più

OSSERVAZIONE

Il coeff. che moltiplica $Y[n]$ è sempre = 1 \rightarrow cioè il 1° coeff. di z è sempre = 1.
 \rightarrow Il vettore a deve essere denormalizzato al 1° elemento, che è quello che moltiplicherà Y

NOTAZIONE

• Se: $mb = \alpha \Rightarrow$ le fdt diventa una costante polinomio \rightarrow le uniche radici sono (POLI)

il tipo di il filtro IIR è: All pass, ricorsivo, autoregressivo (AR). Vol dire che, se ho ad es.:

$H(z) = \frac{1}{1 + 3z^{-3}}$, con $mb = \alpha$. Questo filtro, al denum., è di ordine

Alora:

Il filtro ha 3 poli ed è AR \rightarrow cioè vol dire l'uscita ricorre su se stessa, cioè servono più valori dell'uscita, per creare l'uscita al passo attuale. Cioè $Y[n]$ non dipende solo da X , ma anche da Y ai campioni precedenti.

\Rightarrow quando si intende filtro ricorsivo si intende che l'uscita ricorre su se stessa; vogliamo campioni precedenti dell'uscita, per generare la nuova uscita al campione attuale.

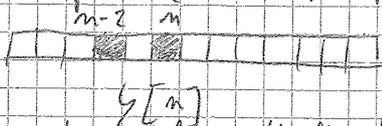
Si dimostra che:

Quando le fdt. è fatta solo da poli, il filtro è IIR, che vol dire: risposta all'impulso infinita (tipicamente diventa una fine asintotica).

• Se $mb = \alpha$, il denum. vale 1; quindi se $mb = \alpha$ tutte le radici delle fdt. sono (ZERI) \rightarrow il filtro FIR è "all-zero", non ricorsivo, è media mobile (MA). Cosa vol dire a media mobile?

Se ho: $H(z) = \frac{2 + 3z^{-2}}{1} \Rightarrow Y[n] = 2X[n] + 3X[n-2]$

A media mobile perché l'out è una combinazione lineare dei un n prefissato di ingressi. In realtà le cose sono molto semplici, perché se voglio implementare un filtro di questo tipo, mi occorre che per generare l'uscita Y , mi serve un campione dell'ingresso (che moltiplico per 2) e mi servono un campione dell'ingresso due passi prima.



\leftarrow mi bastano solo quei 2 campioni lì per generare $Y[n]$, non serve tutto il segnale.

\Rightarrow A media mobile perché: prendo solo i campioni che mi interessano che al passare di n , si spostano lungo gli

\rightarrow L'ordine del filtro è data quanti elementi dell'In e dell'out ci servono per effettuare il filtraggio (al max 19/12 e 12) e magari ci sono anche su lungo da filtrare \rightarrow non importa, si prendono tutti i campioni che servono e non meno si va avanti con il filtro

Un filtro con $mb = \alpha$, si dice FIR, con risposta all'impulso finita.

filtraggio all'input esiste sempre. Allora questo fa capire che conviene lavorare con un ordine del filtro il più basso possibile, perché:

- 1) ho meno operazioni da fare (dato che ho meno coeff.) → però questo mi disturba meno perché sono operazioni semplici;
- 2) se l'ordine del filtro è basso ⇒ il transitorio di filtraggio iniziale, di conseguenza, è basso → questo è il motivo principale!

Però avanti si vede che \exists una regola pratica che dice che l'ordine del filtro non deve mai superare $\frac{1}{3}$ della lunghezza del segnale da filtrare → questa regola è di buon senso non è una regola matematica. Ma nel nostro caso non è un problema più grave di tanto, perché anche se a volte i segnali biologici sono corti, però le frequenze di campionamento sono sempre di ~ 100 Hz ⇒ basta 1 secondo di segnale, per avere già 100 campioni e normalmente ce le copriamo con filtri che hanno ~ 10 -15 ordine al massimo. Quindi il motivo per cui si sceglie che bisogna mantenere l'ordine del filtro il più basso possibile, deriva dal fatto che all'aumentare dell'ordine, aumenta il transitorio del filtro digitale.

Per verificare ancora meglio cosa succede, vediamo come è fatta la struttura ^(SIOE) e minimo numero di ritardatori: si è detto che un filtro digitale è fatto da moltiplicatori (perché bisogna moltiplicare per i coeff. a e b), sommatore (perché è un plusone) e ritardatori discreti di k -passi.

Allora, della sequenza $H(z)$, si può disegnare quello che è la struttura a minimo numero di ritardatori (SIOE S), cioè ~~è~~ è uno schema a blocchi della funzione di trasferimento, che è fatto solo da moltiplicatori, somme e ritardatori discreti di 1 passo in cascata.

Si può vedere che: $\boxed{\frac{z^{-1}}{z}}$ sono tutti ritardatori discreti di 1 passo e ce ne servono tanti, quanto è l'ordine del filtro. Se devo arrivare fino ad mb o me , ce ne devono essere tanti almeno quanto sono mb ed me .

Qui ci sono dei sommatore \oplus , che sono quelli che eseguono la somma dei vari pezzi (quindi dei vari ritardatori in cascata) e poi ci sono i moltiplicatori che sono rappresentati così \rightarrow (non sono frusta, bisognerebbe farli così \rightarrow),

che sono amplificatori \times , perché moltiplicano per il coeff. opportuno (a o b che sia). Quindi in realtà sono \rightarrow moltiplicatori, somme e ritardatori discreti di 1 passo. Vedendo come è fatto questo filtro alla slide 9, le parti di sopra sono tutti coeff. b , che sono di ordine decrescente (da bx verso bx^k): si inizia da quello di ordine più elevato, e si finisce con quello di ordine più basso, $b(1)$

TRANSITORI IN FILTRAZIONE.

Se si suppone di mettere all'ingresso x_1 , si vede come si propaga x_1 all'interno della struttura e così via e poi si mette x_2 e si sposta tutto avanti di un passo e si vede come si propaga nella struttura, si può descrivere il funzionamento del filtro in termini di funzionamento ricorsivo, cioè: al primo passo l'uscita è uguale a $\Rightarrow b_1 \cdot (l'ingresso) + (peso z_1)$ e non c'è altro (perché all'inizio il filtro è silenzioso, ma a regime quando tutto sarà caricato, ovviamente l'uscita è uguale a $b_1 \cdot x_1 +$ quello che esce dal ritardo, cioè: $z_1(m-1) \rightarrow$ ritardo di un passo. $z_1(m-1)$ a sua volta è uguale a: $b_2 \cdot x_1 - a_2 \cdot y_1$ e così via... \Rightarrow si può scrivere l'espressione di z_1 in funzione di quel filtro e quindi si può andare avanti e scrivere il funzionamento di quel filtro in modo ricorsivo.

Allora (S40E10)

Questo è la prima equazione: $y(m) = b(1)x(m) + z_1(m-1)$;
 Poi bisogna esprimere il valore di z_1 : $z_1(m) = b(2)x(m) + z_2(m-1) - a(2)y(m)$;
 Poi z_2, z_3, \dots così via fino all'ultimo z che è uguale a:

$$z_{m-1}(m) = b(m)x(m) - a(m)y(m)$$

Questo non è altro che quello letto come due componenti. È una funzione MATLAB che fa questo (al posto nostro) e lo fa esattamente così, utilizzando (indipendentemente da m_a e m_b) un'equazione per dare l'uscita (che è la prima) e poi non fa altro che usare ingressi e uscite per approssimare i valori z , i quali non sono altro che quelli che troviamo all'inizio dei ritardatori discreti di n -passo $[z^n]$ \Rightarrow è un funzionamento ricorsivo, del tipo: "approssimo l'uscita non meno che ci mette degli ingressi che questi valori z si propagano in avanti (dove in avanti, nella struttura, vuol dire verso sinistra)".

Perché si introduce un ricorsivo?

Di fatto queste serie di equazioni ci dice che il transitorio si finisce perché i valori z all'inizio sono nulli; diventano $\neq 0$ (assumendo il valore corretto), dopo avere messo almeno m (m_a o m_b o secondo di quello che è maggiore tra i due) ingressi e uscite. Solo allora, tutti i valori z hanno il valore corretto e iniziano a propagarsi tutti verso l'uscita.

Fino a quel momento, il m è sempre qualcuno che è nullo \rightarrow si può infatti dimostrare che fino a che ho messo in ingresso un numero ^{di campioni} inferiore all'ordine del filtro, c'è qualche valore z che è $\neq 0$, e che non è caricato completamente.

Il mio punto dove interesso lo f_c , perché per effettuare tecnicamente l'operazione di filtraggio, lo f_c non serve a niente (proprio perché MATLAB segue lo schema della struttura a min. num. di rit. e ragiona in campioni \rightarrow ad ogni ritard, così ad ogni campione successivo, da in uscita qualcosa di più).

Accoppiato alla funzione `filter`, esiste un'altra funzione che ci salva la vita, che è:

`freqz(b, a, 256, 2000)`

\swarrow no di punti
 \swarrow freq. di campionamento

È la funzione che fa il passaggio dalla trasformata Z , alla trasformata di Fourier \rightarrow cioè ci riporta alle funz. di trasferimento da $H(z)$ ad $H(\omega)$ (che è quella che ci interessa).

Questa funzione `freqz` consente, dati i soliti vettori b e a , di rappresentare modulo e fase della funzione di trasferimento, cioè di disegnare il diagramma di Bode.

b ed a dicono dove sono le radici (dicono come sono fatti i polinomi) e MATLAB disegna diagramma di Bode del modulo e della fase \rightarrow rappresenta in due grafici il modulo della funz. di transf. e la fase della funz. di trasferimento di quel filtro.

Normalmente, quando si parla di frequenza, ci interessa avere l'asse orizzontale delle frequenze che non sia normalizzato, ma che sia DENORMALIZZATO \rightarrow cioè se voglio essere sicuro di aver implementato in modo corretto un filtro P-BASSO che taglia a 100 Hz, la freq. di taglio che vedo da quando il modulo della funz. di transf. inizia a scendere bisogna vedere a cosa corrisponde \Rightarrow (alle slide 12, in figura taglia a ~ 400 Hz) e quindi dire avere un'asse orizzontale fornito in Hz, non in unità normalizzate, perché altrimenti si fa fatica a capire. Allora, alla funzione `freqz`, si può dire anche la frequenza di campionamento (che è il quarto parametro).

Il terzo parametro della funzione `freqz`, è il no di punti su cui `freqz` rappresenta modulo e fase. Bisogna dire anche il no di punti perché la transf. di Fourier è una quantità continua in frequenza e quindi si può rappresentare su quanti punti si pare e MATLAB non lo sa \Rightarrow prenderebbe due punti e così se non gli viene detto (di default prende 512 punti). La scelta di questo parametro è una!

BEPER CLASSICO (cioè quando si fa filtraggio, che quando si parla lo mostra del filtro): cambiare b con a (cioè cambio di poli e zeri). L'ordine di tutte le funzioni MATLAB è sempre: prima la parte a medio indole, e poi la parte autoriservata.

NUMA (2010-10)

- Normalmente, in MATLAB, il modulo delle funzioni di trasferimento è fornito in dB.
- a e b sono i vettori dei coeff.
- Un filtro di Butterworth restituisce un filtro passa banda (ha sia poli che zeri) dello stesso ordine (n). Il n° di coeff. è sempre $1 +$ dell'ordine del polinomio (in quanto sono suoi zeri).
- Questo filtro in figura non è a fase lineare, ha un andamento che sembra una sigmoide (questa è una caratteristica comune a tutti i filtri APOL, ~~in generale~~ in generale per i filtri IIR \rightarrow la fase non è mai lineare)

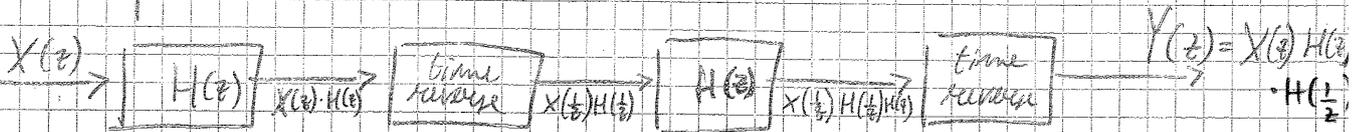
All'interno di Matlab ci sono 2 insiemi di funzioni differenti che sembrano di generare i parametri a e b , per filtri di tipo IIR e filtri di tipo FIR. Sono 2 insiemi di funzioni separate, perché i filtri IIR comprendono sia poli antrospassivi puri (solo poli), sia tutti quelli che hanno poli e zeri; mentre i filtri FIR hanno solo zeri. Allora per comodità di conversione numerica, è più facile tenere le strutture separate; però vediamo che alla fine noi in base, perché tanto (indipendentemente da AP, PA, APOL) se guardiamo le caratteristiche dei filtri FIR e le caratteristiche dei filtri IIR, si vede che le caratteristiche sono un po' complementari e quindi a volte ci serviva generare un filtro FIR e a volte ci serviva generare un filtro IIR.

Vietato filtrare un sgn ECG con filtri a fase non lineari! Perché altrimenti si distorcerebbero le armoniche e conseguentemente l'intero il sgn . Se invece il sgn è CASUALE (e quindi cioè le armoniche sono combinate casualmente), se anche distorco la fase non ^{ne} fa nulla. Il sgn casuale, per definizione, non porta tanto info. nel dominio del tempo (anzi per niente), in la ricostruzione temporale è una concolita \Rightarrow in questo caso filtrare il sgn anche con filtri a fase non lineari, è lecito. Come vuol dire? Ho dei filtri molto presentanti (perché su sgn con piccoli occhio, hanno delle buone prestazioni), ma che su certi sgn non fanno cose, perché non hanno una fase lineare. Come faccio?

Un trucco per recuperare la distor. di fase: filtraggio out-of-phase a scotatura di fase nulla (filtraggio a doppia passata) - (slide 35)

Questo tipo di filtro funziona in questo modo:
 Supponiamo di avere un sgn da filtrare di cui non vogliamo distorcere la fase che è la sequenza numerica discreta $X[n]$, la cui trasformata z è: $X(z)$ e supponiamo di avere un filtro digitale definito dalla sua funz. di transf. $H(z)$, a è un filtro IIR (cioè un filtro a fase non lineare).
 Il filtraggio a doppia passata prevede di sfruttare una proprietà della trasformata z per cui filtrando il sgn 2 volte, dopo un operazione di time reversal (cioè invertito il verso dell'asse temporale), mi consente di recuperare completamente la distorsione di fase.

Nel dettaglio:



All'uscita del 1° filtro ho: $X(z) \cdot H(z)$ \rightarrow In questo momento ho distorto il sgn , però $H(z)$ ha una fase non lineare \Rightarrow all'uscita del 1° blocco, la fase è già andata. Se ci fermassimo qui, avremmo filtrato il sgn P-BASSO, ma vedremmo nel dominio del tempo una distorsione del sgn . Allora a questo punto prendiamo il sgn filtrato e applichiamo un'operazione di time reversal \rightarrow cioè vuol dire che il primo campione diventa l'ultimo e l'ultimo diventa il primo \Rightarrow concettualmente facciamo il vettore che contiene il sgn completamente di 180° (matematicamente è come invertire l'asse del tempo). Guardando la def. di trasformata z e anche le def. delle transf. di Fourier, perché le proprietà sono coincidenti, se invertito l'asse del tempo è come mettere un (-1) all'esponente, in questo caso è come mettere un (-1) e $z \Rightarrow$ l'inversione temporale sarebbe come moltiplicare lo stesso quantito, non funzione di z una funzione di z alle $(-z)$ (o di $\frac{1}{z}$). Quindi dall'operazione di time reversal si ottiene: $X(\frac{1}{z})H(\frac{1}{z})$.

A questo punto rinfiltriamo il sgn ancora una seconda volta con lo stesso filtro

Se invece, ad es., ci viene detto che il filtraggio (Non il filtro) deve introdurre un'attenuazione complessiva di 100 dB \Rightarrow si può scegliere se fare un filtro che attenua 100 dB e applicarlo una volta sola, o un filtro che attenua 50 e applicarlo a doppio passato.

QUINDI: FILTRO E FILTRAGGIO SONO 2 COSE DIVERSE!
(Bisogna leggere bene le specifiche)

Se invece ci viene detto che il filtro deve avere fase lineare o non deve distorcere il segnale \Rightarrow ci stiamo riferendo al filtro e in questo caso bisogna utilizzare obbligatoriamente un filtro FIR (perché un IIR ha sempre fase non lineare e quindi non si può avere).

Se invece il filtraggio non deve introdurre distorsione di fase, si può avere invece un IIR a doppio passato.

Quindi un conto è quello che deve fare la macchina del filtro, un conto è l'operazione di filtraggio (e si può usare tutto quello che si ritiene più corretto, perché per domande di questo tipo non c'è una soluzione unica, ci sono più soluzioni corrette).

② Il filtraggio a doppio passato è nato soprattutto per recuperare la distorsione di fase e quindi si usa solo con filtri IIR.
(Non usare con filtri FIR! Sono leciti solo per filtri IIR e per questi è un'inevitabile distorsione, la fase, altrimenti il segnale si filtra normalmente)

SVANTAGGI

Tra gli svantaggi che ci sono, c'è certamente il fatto che si perde una cosa che dal punto di vista dei filtri digitali può essere una necessità: c'è la possibilità di funzionare in tempo reale. Un filtro applicato a doppio passato, ovviamente, non può essere applicato in real time perché bisogna avere tutta la sequenza per poterla invertire (non si può fare compressione per compressione).

Il bello visto per la struttura o minimo numero di ritardatori, è proprio multimediam campione in ingresso, viene detto come il prodotto il campione in uscita in tempo reale \rightarrow ogni volta che entra un campione, si sbloccano i ritardatori e esce un campione dell'uscita.

\Rightarrow Tutti i filtri digitali (FIR, IIR, ecc...) sono perfettamente implementabili in tempo reale, se però si vogliono fare a doppio passato, No! però deve avere acquisito già tutta la sequenza (cioè bisogna fare un'immagine di tutto).

dovrebbe essere una sinusoide reale \Rightarrow quella è la distorsione di fase.
 Stesso filtro implementato a doppio posto, ha l'uscita che è quell'altra
 sinusoide, che risuona in fase rispetto all'ingresso, perché con la doppia posta
 ho rimosso perduto la distorsione introdotta con il primo filtraggio.

L'istruzione per effettuare il filtraggio si chiama: $\text{filter}(b, a, x)$
 \uparrow
siga da filtrare

L'istruzione per fare il filtraggio a doppio posto si chiama:
 $\text{filtfilt}(b, a, x)$

conclusione: Se ho un filtraggio a doppio posto solo con filter IIR e
 solo quando il compito è non distorcere la fase e quando
 si ha un segnale che non voglio distorcere la fase (limitato)
 perché ad es. può essere un segnale vocale, cioè un che ha una
 forma che si vuole mantenere nel dominio del tempo.
 Per il resto (dove si può...) si filtra normalmente.

SLIDE (17) \rightarrow elenco delle funzioni KARMA (del 1985) che riassumono quali
 sono le funzioni che si possono usare per generare i vettori dei coeff.
 a e b , impone le condizioni del filtro, cioè impone le caratteristiche del
 filtro che si vuole.

Allora:

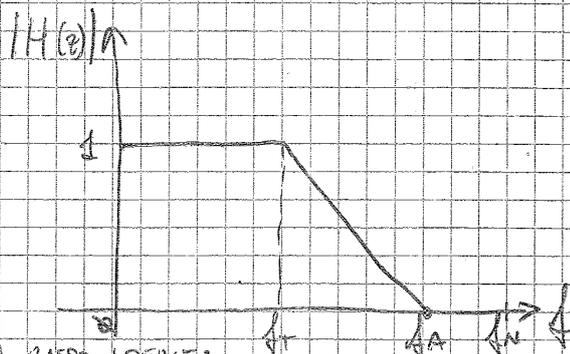
Per filter IIR, ci sono 5 filtri disponibili:

- ✓ bessel (filtro con approssimazioni di Bessel)
- ✓ butter (" " " " Butterworth) (oscillazioni)
- ✓ cheby1 (" " " " Chebyshev con ripple in banda passante o in
- ✓ cheby2 (" " " " Chebyshev con ripple " " " "
- ✓ ellip (filtri ellittici)

Le sintesi sono tutte uguali, vogliono l'ordine e la f_T : la f_T è o
per P-BASSO e P-ALTO, sono due per P-BANDA e RIGETTA BANDA.

d'ordine di un filtro invece per stabilirlo non esiste una funzione
 matematica che si può applicare a priori per sapere quale è l'ordine di
 un filtro di ... che taglia a ... Hz e che a ... abbia già 20dB
 \rightarrow è troppo complicato e quindi detto che per generare i coeff. a e b

questo metodo bisogna fidarsi fino ad un certo punto).
 Si possono di avere un filtro P-BASSO fatto così:



VOGLIO UN PASSA-BASSO IDEALE:

Ciò che volevo fino alla f_T , dalla f_T in poi sostanzialmente vale 0 (questo è il P-BASSO ideale) \Rightarrow NON ESISTE, perché è fisicamente non realizzabile. Allora ci orientiamo di quello disegnato che non ha un fronte verticale, ma ha un fronte che va più ad una certa quantità \Rightarrow le frequenze che ci interessano per descrivere questo maschero, sono:

- la continua \rightarrow punto vale $|H(e^j\omega)|$ in continua (0)
- la $f_T \rightarrow$ freq. di taglio
- la $f_A \rightarrow$ inizio della banda attenuata
- la $f_N \rightarrow$ fine dell'ore (freq. di Nyquist \rightarrow freq. max)

\Rightarrow Se devo descrivere questo filtro, in modo semplice, posso dire: in continua vale 1, fino alla f_T vale 1, alla f_A vale 0 e alla f_N vale 0. Allora è come avere definito 2 vettori (uno che contiene la frequenza e uno che contiene l'ampiezza delle f di trasf.), che ha esattamente la stessa corrispondenza:

$$f = [0 \quad f_T \quad f_A \quad f_N] / f_N$$

$$M = [1 \quad 1 \quad 0 \quad 0]$$

2 vettori di 4 elementi che definiscono le f di trasf. in modo univoco! Però il vettore delle frequenze, per quanto riguarda MATLAB, deve essere normalizzato tra 0 ed 1 \Rightarrow lo divido per f_N così sono sicuro tutto quello che è nel vettore f , è compreso tra 0 ed 1.

\Rightarrow Succede che l'istituzione di Yule-Walker, prende la maschera disegnata in quel modo, la mette nel sistema lineare e si ricavano i coeff. a e b che posizionano le radici, in modo da avere un filtro IIR che pensavo esattamente questo maschero in cospice. L'istituzione è:

Esempio: genero un filtro FIR con vettore $b = [1 \ 2 \ 3] \rightarrow$ non ho fase lineare, perché non è un vettore numerico simmetrico

Se invece fosse: $b = [1 \ 2 \ 1] \rightarrow$ fase lineare

Oppure: $b = [1 \ 2 \ -1] \rightarrow$ antisimmetrico \Rightarrow a fase lineare

Matlab, quando gli viene chiesto di generare i coeff. di un filtro FIR, tenta sempre di generare un filtro FIR a fase lineare.

SVANTAGGI:

Per soddisfare determinate prestazioni, ci vuole un ordine più alto.

Le tecniche di progetto dei filtri FIR sono 2 quelle più importanti:

- $fix1 \rightarrow$ equivalente di butter dei filtri IIR: vuole l'ordine e la freq. di taglio
- $fix2 \rightarrow$ è l'equivalente di $impinvar$ dei filtri IIR: vuole l'ordine, il vettore delle frequenze e il vettore delle ampiezze e fase ed FIR che rispetta quello mascherato.

Le intenzioni sono:

$$b = \text{fix1}(n, Wm, \text{options})$$

$$b = \text{fix2}(n, f, m, \text{options})$$

ESEMPIO: (SLIDE 18) \rightarrow P-BASSO di ordine 10 IIR con approssimazione di Chebyshev che taglia a 0,3. A 0,3 vale ancora 1 e poi va giù in modo molto rigido (a 0,4-0,5 siamo già intorno a 30dB di attenuazione).

L'equivalente filtro FIR è alle (SLIDE 21) \rightarrow P-BASSO di ordine 10 FIR, taglia sempre a 0,3, ma in questo caso l'ordine è insufficiente perché a 0,3, il modulo della funz. di transf. ha già iniziato a scendere e a 0,5 non è nemmeno a 20dB \Rightarrow è molto più lento e pesante di ordine. Se volessimo un filtro P-BASSO che taglia 0,3 e che a 0,5 ha le stesse prestazioni dell'altro, probabilmente bisogna andare su un ordine

16/03/15, Prima di parlare dei filtri per la rimozione dell'interferenza di rete, vediamo meglio la gestione dell'implementazione dei pesi; quando si è visto, infatti, la struttura a minimo numero di ritardatori, si è detto che la struttura è importante perché dà ad ogni campione di ingresso della sequenza X , cosa succede all'uscita, cioè come viene generato il campione d'uscita.

Nelle slide 3 è riportato una struttura flessibile (che sopra ha la parte a medio mobile e sotto quella autoregressiva), però di volta in volta si possono avere i pesi che non mi interessano e occorre la struttura specifica per il mio filtro.

Prima di ogni ritardatore, ci sono dei valori chiamati z , che derivano dal prodotto dei coeff. a medio mobile per gli ingressi sommato al prodotto dei coeff. autoregressivi per le uscite. Questi valori z , però, non vanno all'uscita perché devono passare tutto una catena di ritardatori che devono passare.

Si era anche detto che questi valori di z sono quelli responsabili del transitorio perché all'inizio appena il filtro è stato costruito ma non è stato inserito ancora nessun campione X all'ingresso, ~~possono essere~~ ~~quelli a~~ ~~quelli valori~~ di z (ma non è importante), ma fatto sta che non sono il valore corretto, ovvero non sono b. ingresso e uscita e questo determina il fatto che finché almeno l'ultimo z non si è propagato all'uscita, questa catena di ritardatori genera il transitorio.

Il transitorio di filtraggio non è un problema grave, perché abbiamo detto che avremo sempre un ritardo che è \ll rispetto al numero di campioni che costituiscono il sgn , \Rightarrow avremo transitori di lunghezza abbastanza breve che non ci comportano problemi sui sgn .

È però un modo di pre-inizializzare quei pesi e in alcuni casi può essere importante per annullare o minimizzare il transitorio di filtraggio, in particolare succede quando devo effettuare quello che è detto: filtraggio a blocchi ^(slide 13) \rightarrow il filtraggio a blocchi è una tecnica importante che si usa quando non si può allocare nel mio dispositivo memoria sufficiente per filtrare tutto il sgn \Rightarrow non è tutto il caso del calcolatore con MATLAB, ma è più il caso di dispositivi reali processabili, dove la RAM è limitata e non si ha spazio e suff. per allocare tutto.

Quindi se ho un sgn abbastanza lungo, non lo carico in memoria e lo filtro tutto in una volta, ma lo faccio a pezzetti, cioè lo divido in blocchi. Se ragiono in termini di filtraggio tradizionale, succede che: prendo un blocco del sgn , ~~quello~~ filtro questo blocco, ho un transitorio di filtraggio all'inizio, quando ho finito di filtrare questo blocco, lo cerco, prendo!

del valore z alla fine del filtraggio del primo blocco, diventano le condizioni iniziali all'inizio del filtraggio del secondo.

$$Y_2 = \text{filter}(b, a, x_2, z_f)$$

dove:

Y_1 e Y_2 contengono i 2 segnali filtrati che unisce uno in coda all'altro, per ottenere il segnale filtrato complessivo.

Questo metodo di filtraggio a blocchi (e quindi di pre-correzione di fase), vale solo se si ha 1 solo segnale che si decide di filtrare o perzetti, perché pre-correzione di fase che decidono se un altro segnale, il filtraggio di un altro segnale, non ha senso; invece se il segnale è sempre lo stesso, funziona ed evita il ritardo di fase all'inizio del filtraggio di $Y_2 \Rightarrow$ è una tecnica buona.

NOTA

z_i e z_f sono vettori perché a i dentro un numero di valori, pari al numero di ritardatori della catena \Rightarrow tanti sono i ritardatori, tanti sono i numeri z .

In che rapporto sta la lunghezza di z con l'ordine del filtro? si ripete in MATLAB

Questa tecnica di filtraggio a blocchi vale per qualunque filtro digitale \Rightarrow una volta che si hanno i valori a e b e si ha il segnale da filtrare, qualunque filtro può essere implementato con filtraggio a cascata o filtraggio a blocchi (non cambia niente).

RIMOZIONE DELL'INTERFERENZA DI RETE

Sono filtri che servono per rimuovere l'interferenza di rete che per i segnali biomedici, è particolarmente importante perché le nostre condizioni sperimentali fanno sì che nonostante noi tentiamo di limitare o prevenire la presenza dell'interferenza di rete nel segnale che misuriamo, spesso le cose non vanno come vorremmo e quindi il fatto che ci sia un po' di interferenza di rete nell'acquisizione che facciamo, lo dobbiamo tenere in conto.

I segnali biologici che risentono di più dell'interferenza di rete sono quelli che hanno un'informazione specifica nel dominio del tempo, che voglio mantenere perché la 50 Hz si sovrappone e la altera \Rightarrow il segnale che forse più di tutti patisce la presenza di 50 Hz è di certo il segnale ECG. Ci sono altri segnali che vengono misurati con una banda un po' più grande e un volta anche:

In generale, se $f_c = k \cdot f_0$, allora un filtro notch avrà l'espressione

$$y[n] = x[n] - x[n-k]$$

periodo di ripetizione \rightarrow f_0

Dall'equazione temporale di un filtro, abbiamo invece in grado di passare subito a ricavare la funzione di trasferimento $H(z)$ e quindi si scrive subito

$$H(z) = 1 - z^{-k}$$

Questa è la funzione di trasferimento di un FIR \rightarrow il grande vantaggio di questo filtro notch è che, oltre ad essere un filtro FIR, è di conseguenza SEMPRE STABILE per qualunque valore di k e può avere fase lineare \rightarrow C.N.E.S. di un filtro FIR abbia fase lineare è che il vettore dei coeff. b abbia elementi numerici o antisimmetrici.

Per la matrice $H(z)$, abbiamo che il vettore b è:

$$b = [1 \underbrace{0 \ 0 \ 0 \ \dots \ 0}_{z^{-1}} - 1] \rightarrow \text{è un vettore antisimmetrico} \Rightarrow$$

Questo filtro FIR è a fase lineare \Rightarrow questo è il bello del filtro notch soprattutto in termini di ECG: è SEMPRE STABILE e a FASE LINEARE \Rightarrow non distorce il segnale. Questa condizione di non distorsione è importante nel caso di segnali che portano info nel dominio del tempo.

Inoltre, essendo un filtro FIR canonico anche il vettore a : $a = 1$.
 \Rightarrow un filtro notch è un filtro equifase e se bisogna implementarlo su un microprocessore si sceglie perché conosco sia la f_c , sia quanti e quali sono gli elementi e quindi si bandisce per il filtraggio.

NOTA

Può capitare che f_c e f_0 non siano multipli interi, cioè non ci sia un valore di k intero tale $\frac{f_c}{f_0}$ (caso più comune: invece di avere $f_c = 250$, si usa la potenza del 2 più vicina: $f_c = 256$ in questo caso). In questo caso non devo assolutamente prendere il valore di k più vicino, ma devo ricompilare (perché deve essere assolutamente i campioni nella stessa posizione periodo per periodo) ad una f_c opportuna (multiplo della f_0) e si può fare, grazie al th. di Nyquist che dice che se ho rispettato con la f_c il teo. il segnale è perfettamente ricostruibile \Rightarrow l'ho acquisito, lo ricostruisco lo ricompilano alla frequenza che mi pare. Forse ciò teoricamente non è un problema, il nemmeno tecnicamente perché esiste una funzione che

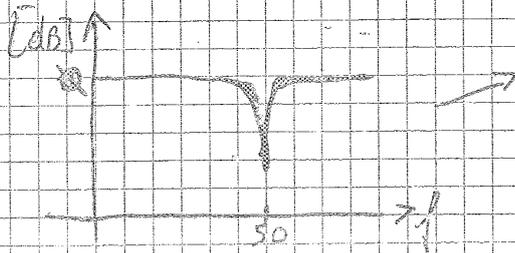
oppure che voglio una $f_c = 250 \text{ Hz}$ (e partendo da una f_c maggiore), applico un filtro P-BASSO ^{che tagli} a 125 Hz e poi ricampiono a 250 Hz il segnale in modo da evitare il più possibile il fenomeno dell'aliasing.

ATTENZIONE

Il fatto di non avere aliasing perché ho filtrato P-BASSO, non sempre va bene se sottocampionando ho buttato nella spazzatura metà del segnale, non va bene \Rightarrow bisogna stare attenti a come funzionano questi strumenti e poi come li usiamo.

Se uso l'intenzione di fare per esempio a vedere come è fatto questo filtro in modulo e fase, scopro queste cose qua: 54:05:33

L'unica cosa che va bene è la fase, che effettivamente è lineare; ma guardando il modulo, mi accorgo di almeno 2 cose che non mi piacciono:
 1) il modulo delle f di transf. non ha ~~un~~ valore costante ed è, in bande passante, infatti abbiamo detto che teoricamente noi vorremmo dei filtri che in bande passante hanno un valore pari a 1 (cioè 0 dB) con il segnale non lo tocchiamo. La banda passante, in questo caso, sarebbe fatta ad eccezione di 50 Hz (che deve essere attenuata) \Rightarrow in questo caso non c'è un modulo costante pari ad 1, ma c'è il fenomeno dell'OVERSHOOT, cioè la funz. di transf. non è la f_c ed è superiore ad 1, ma poi come sono posizionati gli zeri, da origine ad una forma delle funz. di transf. che ha un valore che supera il valore unitario e poi torna giù \rightarrow il motivo per cui succede questo è che non ci sono abbastanza radici per dare tempo, perché noi vorremmo una funz. di transf. così:



Per riuscire ad avere questo comportamento, i poli devono essere molto ripidi e per a ciò, ci vogliono più radici. In un filtro non quello che determina il numero di radici è $k \rightarrow$ ma $k = \frac{f_c}{f_0}$ se fissa e quindi non

si può mettere a caso, non lo so io ma dipende da f_c e f_0 \Rightarrow il numero di radici è fisso, perciò avendone così poche e dovendo far così tante cose, la funz. di transf. non è la f_c e crea questo overshoot grande passante (normalmente non è un overshoot molto elevato, ci va bene...)

il pa y di uscita è sempre un po' a valore medio nullo \Rightarrow nell'ordine
 della \sim me 100 volte questo po' a po' nell'uscita è un filtro che
 mi consente di avere po' un po' in uscita a valore medio nullo e se poi loro
 dopo essere filtrate, ci voglio fare una spettro, non ho problemi di valore
 medio.

Dato la funzione di trasferimento $H(z) = 1 - z^{-K}$, la sua f di f
 $\hat{H}(f) = 1 - e^{-j2\pi fK}$, dove: $H(f)|_{f=0} = 0$, \forall valore di K.

\Rightarrow è facile dimostrare che il filtro notch ha uno zero di trasmissione in
 continuo.

Inoltre il filtro notch funziona in real-time nel senso che è ovvio che
 devo avere almeno K campioni precedenti, ma finché non ho i K campioni è
 il transitorio di filtraggio; ma ^{non} appena sciro il campione, sono pronto a
 fluire l'uscita senza problemi.

2° FILTRO: FILTRO RICORSIVO (SLIDE 33)

Se non funziona il filtro notch, funziona di certo il Filtro Ricorsivo.
 Un filtro notch non funziona quando le ip fondamentali sono ritate, cioè
 quando l'interferenza di rete non ha frequenza, ma ^{non ha} ~~è~~ ~~costante~~
 costante! In questo caso si fa un filtro ripple-banda tradizionale, chiamato
 Filtro Ricorsivo (perché è un IIR) ma di fatto è quello noto come
 ripple-banda tradizionale \Rightarrow ha una funz. di transf. come in SLIDE 36
 il modulo è 1 ovunque, tranne a 50 Hz dove introduce una certa attenuazione
 invece la fase è un disastro perché essendo un filtro IIR, non ha fase lineare con
 tutti i filtri IIR \Rightarrow questo filtro effettua un buon filtraggio anche in caso di
 ampiezza non costante della sinusoidale di rete, ma non è implementabile in
 real-time e quindi è senza distorsione \Rightarrow questo è il difetto principale di
 questo filtro.

È moltissimi tipi diversi di ripple-banda, ma quello che consideriamo noi è il
 ripple-banda particolare fatto da 4 radici: una coppia di poli e zeri complessi
 coniugati (SLIDE 33) \Rightarrow si sa già che è un filtro passa banda di ordine 2 \Rightarrow
 so già che i vettori a e b dei coeff. avranno 3 termini ciascuno.
 Allora il problema del filtro ricorsivo è posizionare le radici, per dare
 forma ad una funzione di transf. che sia un ripple-banda. (63)

B : larghezza di banda a 3 dB (cioè la larghezza "dell'imbuto" e unita)

f_0 : la frequenza di risonanza (50 Hz)

T : l'intervallo di campionamento (cioè: $\frac{1}{f_c}$)

In uscita restituisce:

e_{MA} → coeff. a media mobile

e_{AR} → coeff. autoregressivi

Allora:

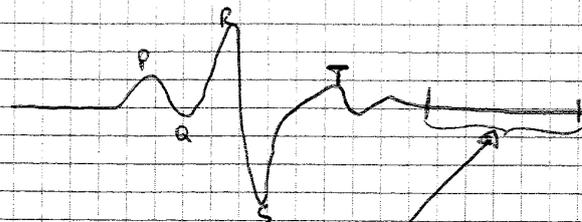
$$[e_{MA}, e_{AR}] = \text{rico}(z, B, f_0, T)$$

Bisogna dare tutti questi parametri in ingresso perché il filtro ricorsivo ha un vincolo: ha 4 radici \Rightarrow è secondo di come le posizioni, non può avere una banda stretta e un'attenuazione (z) molto grande, perché il filtro non va più e più di 20 dB/dec. Quindi questo filtro con n° di radici fisso impone che io devo mettere in modo corretto il valore di z con il valore di B e f_0 e T essi tenterà di posizionare le radici in modo da garantire il valore di attenuazione (z) e la banda (B) ottimali, ma se si fa più che non si può avere obiettivamente z elevato e B piccolo perché il n° di radici è insuff. Allora, se volessi avere una banda molto stretta (cioè un risonanza-banda di risonanza solo la 50 Hz), cioè i poli li devo mettere molto vicini; ma se li metto molto vicini, le funz. di transf. non fa in tempo a scendere, che fa come subito lo zero che la tira su \Rightarrow ottiene poca attenuazione in questo caso.

Se voglio tanta attenuazione, devo allontanare le radici; ma allontanando le radici la funz. di transf. ha il tempo di scendere, ma diventa molto lungo \Rightarrow non voglio solo a 50 Hz, ma voglio qualcosa anche all'intorno di 50 Hz.

Quindi:

La funzione rico serve per posizionare le radici, ma bisogna sempre controllare le radici, perché si rischia di avere un imbuto molto largo e quando serve a filtrare, il segnale filtrato non viene bene, cioè posso avere un ECG in cui il picco Q-R-S l'ho abbattuto, perché ho filtrato più roba e non solo 50 Hz.



cioè, ecco alcuni tratti, tipo questo (in cui c'è la linea isoelettica tra un'onda T e l'onda P del battito successivo), in cui c'è solo $m(t)$ → il spu in realtà $\hat{i} = 0$ ⇒ l'hp di funz. di questo filtro è che il spu, di cui abbiamo fatto denoising, abbia questa caratteristica, cioè: un intervallo di tempo suff. lungo in cui sia presente solo rumore (cioè il spu deve essere nullo o al più costante).
Tutti i spu non intermittenti, ovviamente, non funzionano bene con la tecnica del filtro adattivo.

Si fa questo hp per 2 motivi:

- 1) come detto, nell'intervallo considerato, $x(t) = m(t)$ → si può dire che in quell'intervallo, quello che ho misurato, è solo rumore (ci sono solo componenti di rumore);
- 2) essendo un filtro adattivo, per riuscire ad adattarsi, ci mette un po' di tempo ⇒ quell'intervallo isoelettico dove il spu è costante, è l'intervallo di adattamento del filtro; cioè all'interno di quella finestra temporale, il filtro si adatta e comincia ad inseguire la sinusoide di rumore. Se questo intervallo fosse assente (o fosse troppo stretto), il filtro non funzionerebbe bene.

Quindi:

è sempre un'hp sul spu (e questo non ci piace), però è un'hp ragionevole, almeno per il spu che più di tutti soffre dell'interferenza di rete, che è il spu ECG.

COME FUNZIONA QUESTO FILTRO?

Supponiamo di indicare con:

X_m → i campioni del rumore di rete ~~misole~~

L_m → i campioni (che generiamo noi...) di una sinusoide che ha freq. pari alla freq. che vogliamo ripetere e che noi usiamo per stimare X_m

Y_m → segnale filtrato ⇒ $Y[m] = X[m] - L[m]$

Allora:
$$l_{m+1} = 2k l_m - l_{m-1}$$

Ciò: dato un campione attuale (l_m) e dato il campione precedente (l_{m-1}), può cercare di predire l'errore che si commette al passo futuro (cioè al campione successivo l_{m+1}).

Se il filtro adattivo funziona bene (cioè cerca e predice in modo corretto il futuro), succede che quando scriverò il nuovo campione reale, il valore di questo nuovo campione reale sarà uguale a quello che ho predetto, cioè:

$$l_{m+1} = x_{m+1} \rightarrow \text{in questo caso sono in situazione di ADATTAMENTO, cioè}$$

dato che devo inseguire una sinusoidale, ho messo e posto tutte le ampiezze della sinusoidale in modo che effettivamente sto prediccando il campione futuro sulle basi dei campioni requisiti.

Se invece la predizione non è corretta, ovviamente esiste un errore di predizione e per sapere di quanto ho sbagliato, si fa la differenza tra quello che è e quello che ho detto io, cioè:

$$d_{m+1} = (x_{m+1} - l_{m+1}) - (x_m - l_m)$$

Questa differenza, chiamata anche distanza, è appunto una distanza progressiva che misura l'errore di precisione al passo attuale e l'errore di precisione al passo futuro, in modo tale da seguire in estrema tutti gli errori di predizione che compio e quindi man mano tento di adattarmi alle stesse.

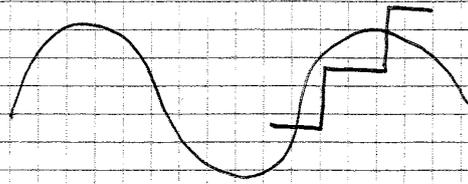
Se la stima è corretta, non si fa niente; se invece la predizione non è corretta ci sono 2 casi:

• Se: $d_{m+1} > \alpha \Rightarrow$ sto sottostimando \Rightarrow aumento il valore del campione futuro, cioè: $l'_{m+1} = l_{m+1} + c$

• Se: $d_{m+1} < \alpha \Rightarrow$ sto sovrastimando \Rightarrow diminuisco il valore del campione futuro, cioè: $l'_{m+1} = l_{m+1} - c$

dove: c è una quantità arbitraria, ma ragionevolmente piccola in proporzione a quella che è d_{m+1} e questa quantità c è una costante \Rightarrow ad ogni passo la quantità che si aggiunge o sottrae è sempre la stessa.

2) C non può essere nemmeno una quantità troppo grande, perché altrimenti la sinusoidale verrebbe smorzata così (e prodotta):



Ciò si continua ad oscillare intorno alla sinusoidale di una quantità molto grande e anche questo non è una buona stima.

Allora:

Ottimizzare il valore di C è un problema, nel senso che non si va a caso, ma nel caso dell'ECG la frequenza cardiaca ci dà l'informazione sul valore numerico di C .

Se la frequenza cardiaca è alta, C dovrà essere più grande perché il tratto isoelettrico è corto e quindi ho poco tempo per adottarmi, rischio di non adottarmi benissimo perché se C è grande c'è questo rischio, ma è l'unico metodo per tentare di adottarmi.

Se invece la frequenza cardiaca è bassa, il tratto isoelettrico è più lungo e mi posso permettere valori di C più bassi.

Una volta che in un tratto isoelettrico questo filtro si è adottato, non è finita qui perché non va tutto bene perché il filtro adattivo è stupido perché fa solo dei paraggi matematici in esterno \Rightarrow il problema nasce quando all'interno del nostro campione x_{n+1} è vero che stiamo paragonando rumore vero e rumore generato solo quando siamo nel tratto isoelettrico ma quando arrivano le onde, x_{n+1} non contiene più solo il rumore, ma contiene segnale e rumore \Rightarrow questo filtro in questo caso si disadotta perché lui tenta di adottarsi e quello che c'è all'interno di x_{n+1} , ma se all'interno di x_{n+1} non c'è più solo la sinusoidale ma ci sono le onde, inevitabilmente tenta di adottarsi anche a tutte le onde P, Q, R, S, T. Allora questo filtro, se funziona bene, arriva all'onda P in condizioni di adattamento, comincia a filtrare e man mano che filtra, si disadotta \Rightarrow arriva all'onda T che si è un po' disadattato e si deve riadattare nel tratto isoelettrico successivo \Rightarrow tutta la vita di questo filtro è un continuo adotta / disadotta ma funziona perché il disadattamento che il filtro ha durante il battito (ciò durante la presenza delle onde dell'ECG)

STIMA SPETTRALE

I motivi per cui è importante vedere gli spetti di stima spettrale sono: i sig. biologici sono le misure che facciamo di un campo elettrico, magnetico o elettromagnetico di un sistema fisiologico e succede che questi sistemi fisiologici lascino la loro impronta sui sig. non tanto nel dominio del tempo, ma nel dominio delle frequenze.

Allora il motivo per cui si fa stima spettrale è perché il contenuto frequenziale del segnale, per molti segnali, è maggiormente informativo del sistema che li ha generati rispetto a vedere l'acquisizione semplicemente nel dominio del tempo → questo è particolarmente vero per i sig. come cioè per quei sig. biologici che sono oscillabili e processi casuali, dove l'informazione portata nel tempo è, a volte, irrilevante.

ESEMPIO: la gran parte dell'info. del sig. EEG sta nella frequenza; stessa cosa per il sig. EMG. Poi ci sono altri sig. che portano una info. nel dominio del tempo e nel dominio delle frequenze un'info. complementare, cioè fanno vedere un'altra cosa o l'effetto di un altro sistema fisiologico ⇒ la stima spettrale, per noi che ~~fare~~ trattiamo sig. biomedicali, è uno strumento fondamentale.

Forse stima spettrale vuol dire ricavare la stima dello spettro del sig., lo spettro a tds è stato visto come: $|X(f)|^2$, i "regolisti" invece definiscono lo spettro in altro modo e nel nostro caso non è applicabile perché non vedo e disegno lo spettro reale, ma il lo spettro con errori, imprecisioni che di fatto è spazzatura ⇒ tutto funziona perché se vediamo come è stato ricavato lo spettro e partire da un sig. $x(t)$, si vede che compone l'integrale delle transf. di Fourier (o la sommatoria) da $-∞$ e $+∞$. Noi però sig. di lunghezza $∞$ non ne abbiamo lavoriamo su pezzi di sig., lunghi ma non di lunghezza $∞$ ⇒ lavoro a tempo limitato per noi non è buono, perché introduce problematiche di errori: lo spettro ^{vero} di un sig. EEG non lo ha mai visto nessuno, sono state ~~viste~~ viste tante stime, ma lo spettro reale non siamo in grado di calcolarlo senza commettere errori o con avere limitazioni ⇒ da qui il nome STIMA SPETTRALE, perché al meglio quello che si può fare è stimare uno spettro, non certo calcolarlo.

La stima spettrale si divide in 2 famiglie: STIMA SPETTRALE TRADIZIONALE (o NON PARAMETRICA) e STIMA SPETTRALE PARAMETRICA

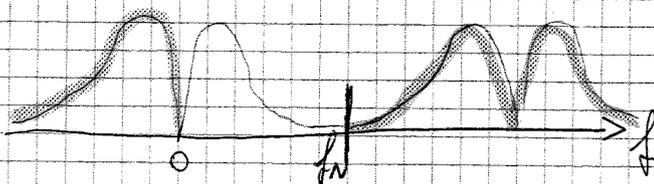
lo $\sum_{m=-\infty}^{\infty}$: noi infatti lavoriamo al finito e quindi noi ci dobbiamo dire non se un $x[n]$ è stazionario in senso lato in generale, ma se quel $x[n]$ può essere considerato un $x[n]$ staz. in senso lato su quel pezzo di $x[n]$ che stiamo considerando.

Ad es. se da un soggetto preleva un $x[n]$ EEG, noi più che non è stazionario, perché il $x[n]$ EEG cambia in base alle funzioni che fa il soggetto \Rightarrow cambia con il funzionamento del cervello. Però ho acquisito, ad es., 1 secondo del $x[n]$ EEG \Rightarrow la voce domanda è se quel secondo è staz. in senso lato. Se sì, in quel secondo di $x[n]$ posso applicare il teorema di Wiener - Khintchine e quindi posso estrarre lo spettro del $x[n]$.

NOTA:

Abbiamo parlato di spettro, ma quello che in realtà andiamo a misurare è uno spettro di potenza (non di energia \rightarrow è un'altra cosa, perché l'energia dipende dal tempo) \Rightarrow invece di dire spettro di energia, otteniamo DENSITA' SPETTRALE DI POTENZA (PSD): è un grafico che fa vedere come la potenza è distribuita nelle frequenze. La $W-K$ della formula di $W-K$ lo spettro di potenza che chiameremo PSD del segnale o più correttamente STIMA della PSD del segnale.

Il teorema di $W-K$ dice che se effettivamente queste ipotesi di $x[n]$ stazionario è rispettata, siamo sicuri che la PSD che abbiamo ottenuto è simmetrica, reale e positiva. • Simmetrica rispetto alle frequenze ω o rispetto alle freq. di Nyquist; ma perché è simmetrica? La transf. di Fourier è simmetrica e $P_x(f)$ è il risultato di una transf. di Fourier \Rightarrow ogni transf. di Fourier è simmetrica: c'è un semiasse positivo e uno negativo.



Ci sono 2 scuole di pensiero: chi dice che la transf. di Fourier è simmetrica rispetto all'origine e chi dice che è simmetrica rispetto alle freq. di Nyquist \Rightarrow NON CAMBIA NIENTE se vedo la simmetria rispetto all'origine (e quindi ho la replica con) o la ^{vevo} simmetria rispetto alle massime freq. che rappresento nel discreto, cioè la f_n e la fine dell'altra parte così, non cambia nulla: è solo una

potenza in funzione della frequenza, il segnale deve essere fermo, la potenza deve essere sempre distribuita allo stesso modo.

⇒ La stazionarietà prescinde dal tempo e ho un pezzo di segnale che considero con proprietà spettrali costanti e allora vado a fare la stessa.

20/3/15

Il th. di W-K è un'osservazione → dice che la PSD ($P_{x,x}(f)$) è la trasformata di Fourier delle sequenze di autocorrelazione!

PSD e spettro (di potenza, non di energia) sono definizioni coincidenti, ma non esse, ma lo spettro lo considero come $1/T^2$ delle trasformate di Fourier e in effetti la definizione di spettro di potenza per processi ergodici è:

$$P_{x,x}(f) = \lim_{M \rightarrow \infty} E \left\{ \frac{1}{(2M+1)T} \left| T \sum_{n=-M}^M x[n] e^{-j2\pi f n T} \right|^2 \right\}$$

Trasf. di Fourier su un tratto di segnale di lunghezza limitata perché la \sum va da $-M$ a M e quindi considero un segnale con $2M+1$ campioni.

⇒ Se ho un processo casuale di cui ho un certo numero di valori temporali, ne calcolo la trasformata di Fourier discreta, ne prendo il modulo quadro, normalizzo rispetto alla durata temporale di questo segnale e ne prendo il valore atteso, dato che è un processo casuale; successivamente lo spettro coincide con il lim di tutto quello che.

In altre parole: modulo quadro della trasformata di Fourier è lo spettro come lo ~~vediamo~~ vediamo anche noi, ma matematicamente è la media statistica del modulo quadro della trasformata di Fourier di un pezzo di segnale, diviso per la sua lunghezza, al tendere all'infinito della lunghezza delle finestre che considero → questa definizione vale sempre per processi WSS ed ergodici (per un segnale biologico l'ergodicità la supponiamo e prassi).

⇒ È una stima dello spettro!

Ci sono 2 soluzioni a partire dallo stesso segnale: quello che mi dia questa definizione per processi ergodici e il th. di W-K ⇒ SONO LA STESSA COSA, perché lo stima spettrale non simmetrico si divide in 2 parti principali: METODI DI STIMA DIRETTI E METODI DI STIMA INDIRETTI

$$\text{rect}[n] = \begin{cases} 1 & , \quad n = m \leq N-1 \\ 0 & , \quad \text{altrove} \end{cases}$$

Quindi, tutto quello che è fuori delle finestre temporale vale 0.

Supponiamo che di $X_0[n]$ voglio calcolare lo spettro \Rightarrow faccio la trasformata di Fourier e mi faccio il $| \cdot |^2$. Però se il spm lo posso scrivere come prodotto di 2 sig. differenti, la trasformata di Fourier di un spm che è il pro di 2 segnali, è la convoluzione della trasformata di Fourier e inoltre la trasformata di f di $\text{rect}[n]$ è un seno digitale, cioè è una funzione oscillante.

Allora succede che ottengo: (SLIDE 17)

$$X_0(f) = X(f) * D_N(f)$$

dove:

$$D_N(f) = T \frac{\sin(\pi f T N)}{\sin(\pi f T)}$$

SINE DIGITALE
KERNEC DI DIRICHLET

Allora, se ho la convoluzione di 2 segnali succede un coseno. Facciamo un esempio semplice: (SLIDE 18) \rightarrow supponiamo che il nostro spm sia una sinusoidale, perché la trasformata di Fourier di una sinusoidale è una o due δ . \Rightarrow supponiamo sinusoidale ^{centrato} a frequenza f_0 (a) \rightarrow se prendo questa sinusoidale e la traccio (c), prendo le due δ in f_0 e $-f_0$ delle figure (b) e faccio le convoluzioni con la funzione ~~sinusoidale~~ \rightarrow fare questa convoluzione è facile perché si prende questa funzione e la centro sulle $\delta \Rightarrow$ prendo l'andamento del seno digitale e lo centro a f_0 e $-f_0$ (d). Però questo è un ~~sinusoidale~~ ^{dirichlet}, perché vedo una comparsa con tante code che si ripetono, non vedo una δ che mi fa dire che è una sinusoidale il mio segnale. E questo è il caso facile, ma se nel segnale avessi due sinusoidali tracciate, ci sono tutti lobi che si sovrappongono a freq. diverse e si sommano e quindi diventa uno spettro difficilmente comparabile e non è nemmeno quello corretto, perché quello che vogliamo è una δ . Allora questo effetto di finestratura introduce inevitabilmente l'impossibilità di stimare lo spettro corretto \rightarrow non si può, perché abbiamo un n° di campioni finito e quindi succede un modello come in figura (e) che è tanto più complicato, quante più frequenze sono contenute nel spm, perché ogni frequenza è una δ e ogni volta che c'è una δ , occorre la funzione di convoluzione e quindi diventa una cosa illecibile.

La potenza è concentrata alla stessa frequenza, mentre nel grafico di $P_{xx}(f)$ c'è una parte di potenza che non è alla frequenza f_0 , ma è sparpagliata dalle oscillazioni che ha questa funzione: $\frac{\sin(x)}{x}$ (SINC).

Questo effetto è un effetto di sparpagliamento della potenza, o PERDITA DI POTENZA detto LEAKAGE DI POTENZA ed è dovuto alle oscillazioni che ha lo spettro di questa finestra \rightarrow oscillazioni che chiamiamo lobi laterali. Cioè diciamo che lo spettro di una finestra è fatto da 2 parti principali: il lobo principale (quello centrale) e una parte che sono i lobi secondari che hanno una certa area associata \rightarrow la somma complessiva fa 1!

Allora, parto dalle condizioni reali in cui tutta la potenza è concentrata alla stessa frequenza, ed una condizione che cerco e stimare dove non tutta la potenza è concentrata alla stessa ~~stessa~~ ^{frequenza}, ma un po' è sparpagliata sulle frequenze vicine, dall'effetto di perdita che hanno i lobi secondari. La differenza è che una finestra a supporto limitato ha il lobo centrale che è analogo alle δ di Dirac, e poi ha tutte le oscillazioni che sono i lobi laterali \Rightarrow la δ di Dirac è come se fosse una finestra senza lobi laterali, ha solo un lobo principale infinitamente stretto \rightarrow non è fisicamente realizzabile la δ , non ci possiamo mai avvicinare. Dobbiamo solo capire come minimizzare il "digiato" nella figura di $P_{xx}(f)$, sapendo da cosa è generato.

Il problema non è banale, perché noi facciamo stime spettrale di un segnale cui so poco e niente e voglio sapere come si comporta lo spettro in frequenza, perché in molti casi, ~~in~~ molti sistemi fisici, facciamo la loro impronta in frequenza \Rightarrow mi serve sapere quanto potenza c'è ad una det. frequenza, quale è la banda del segnale, ecc...

Ma se questo finestratore implicito mi sparpaglia la potenza in giro per le zone delle frequenze, io non sono tanto sicuro di interpretare bene lo spettro di questo segnale o di rappresentarlo bene.

Questo è un effetto grande ho 1 sinusoidi, ma se ne ho di più (oppure ho un segnale a banda larga e quindi ho tantissime sinusoidi) succede un gran casino.

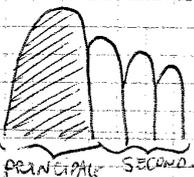
Questo effetto cioè lo chiamiamo LEAKAGE di potenza, più avanti lo chiameremo POLARIZZAZIONE della stima \rightarrow tanto di forza di una stima colgo un valore sbagliato e si dice che la stima è polarizzata, cioè

più basso del lobo primario del caso a) perché nel caso a) l'ampiezza è 1 mentre nel caso b) è 0.19.

Il caso c) è lo spettro della somma dei 2 casi a) e b) → prendo i 2 ^{sinusoidali} \sin li sommo tra di loro e poi faccio la transf. di Fourier → dato che la transf. di Fourier è lineare, allora faccio la somma dei 2 transf. di Fourier e poi prendo il Modulo quadro della somma. Succede che rimane la sinusoide ad ampiezza unitaria centrata a 0.15, e quella che era a 0.24 sparisce totalmente perché il picco dei lobi che si combinano uno sull'altro, di fatto ottiene.

Questo esempio fa vedere come possono ^{essere} ottenute, allo stesso modo possono nascondere componenti cosinusali se i lobi casualmente si sommano ad un certa frequenza. Questo è il motivo per cui il leakage di potenza per noi che poi dobbiamo fare interpretazione, è pericoloso e quindi bisogna tentare di evitarlo → per fare ciò si utilizza una finestra che non è la finestra rettangolare. Quindi invece di prendere il segnale su cui c'è una finestra implicita che vale 1 e calcolare il periodogramma o il correlogramma, posso prendere il \sin e prima di fare la transf. di Fourier, ci metto sopra una finestra che voglio io; ovviamente vorrei la S di D.M.C, ma non la posso fare → devo trovare una finestra che abbia delle caratteristiche comode abbastanza, cioè dato che sono i lobi secondari quelli che danno leakage di potenza, allora voglio una finestra con lobi secondari suff. bassi.

Vedremo finestre con lobi secondari molto bassi, fatte così:



Altra finestra con lobi secondari molto più bassi: 
 Il diagramma mostra un rettangolo con un'area interna hachurata, simile al precedente, ma con lobi secondari molto più bassi e vicini al lobo principale. Sotto il rettangolo sono etichettati "PRINC." e "SEC.".

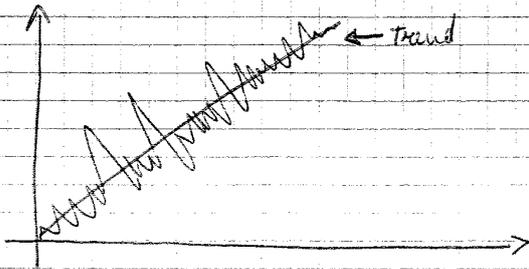
Allora succede che l'area totale, deve valere 1 e se quindi tolgo potenza dai lobi secondari, questa va a finire nel lobo principale, che si allarga → se considero una finestra, l'ampiezza dei lobi secondari ci interessa, ma anche la larghezza del lobo principale ci interessa perché non vogliamo una finestra con lobo principale esagerato. Quindi va bene levare un po' di potenza dai lobi secondari a favore del lobo principale.

perché c'è un potenziale che si sovrappone all'andamento dell'ECG e che una variazione non fisiologica, ma è un potenziale autonomo dovuto ad altre cose.

Allora il trend, se esiste, va rimosso prima di fare stime spettrali non per è particolarmente dannoso, ma perché non ci interessa e che sul piano delle freq. andrebbe a darsi fastidio con il leakage di potenza con tutte le alt. frequenze \Rightarrow lo tolgo prima.

Non si può dare una def. matematica di trend, ma basta un minimo di buon senso. Dato che il trend ~~si~~ ^{si sviluppa} come delle oscillazioni, a volte anche ampie, che si muovono sul ω , se prendo la freq. di queste oscillazioni e la freq. del mio ω , il trend è a freq. più bassa rispetto al ω \Rightarrow per fare detrending uso un filtro PASSA-ALTO alla frequenza giusta (perché se taglia troppo poco il trend rimane, se tagli troppo porta via anche un pezzo di ω).

ATTENZIONE: Ci sono ω (fisiologicamente non tanti...) che sono fatti così:



Se vedo quale è il trend, scopro che è una ω \Rightarrow in questo caso si può fare un detrending polinomiale, cioè prendo quei componenti, faccio fitting lineare con

una retta, ottengo la retta che ci passa in mezzo e la sottraggo e automaticamente ho tolto la deriva rettilinea.

È una funzione in Matlab che consente di fare l'interpol. polinomiale di cui prendo, dato un insieme numerico e si chiama: `polyfit` (FITTING POLINOMIALE) \rightarrow basta dargli il ω , l'ordine del polinomio con cui voglio interpolare (retta ordine 1...) e lui restituisce i coeff. di quel polinomio \Rightarrow posso generare il polinomio e poi lo sottraggo ai componenti e ho fatto detrending.

È un'altra funzione in Matlab che si chiama proprio: `detrending` ma non sappiamo cosa fa! Perché dato che non so def. matematicamente un trend, non so nemmeno costruire una funzione che lo tira via \Rightarrow ~~non~~ si può usare, facendolo funzionare per il mio caso ma non è detto che vada bene... \Rightarrow dato che fare un filtro P-ALTO e fitting polinom. non costa niente, conviene usare `polyfit`.

METODO INDIRETTO \rightarrow CORRELOGRAMMA (Finestra seq di autocorr)
 sulla corte,

Il correlogramma è uguale al periodogramma: acquisizione campioni (un certo numero di campioni e una freq. di comp.), se il caso rimuovo il trend, poi rimuovo il valore medio e a questo punto calcolo la seq di autocorrelazione, così impongo un ritardo massimo tra cui calcolo il valore di correlazione \Rightarrow si vedrà che se di periodogrammi ne esistono tanti diversi, di correlogrammi ne esiste 1! L'unica variabile tra correlogrammi diversi è il numero di ritardi su cui calcolo la seq di autocorr. Poi faccio la transf. di Fourier, vedo il correlogramma se mi piace, ok; altrimenti cambio il ritardo, ci metto le finestre, cambio le finestre, ... faccio quello che voglio e vedo finché il correlogramma non mi va bene.

DIFFERENZA PRATICA TRA CORRELOGRAMMA E PERIODOGRAMMA: la cosa fondamentale di un periodogramma è che è definito come il modulo quadro della transf. di Fourier \Rightarrow per def. è reale e positivo. Mentre nel caso del correlogramma, non è così perché il correlogr. è la transf. di Fourier della funz. di autocorrelazione \Rightarrow è certamente reale (perché autocorr. è pari) ma non positiva, può avere infatti anche parti negative. In teoria non dovrebbe essere così (perché lo dice il th. di W-K), ma il th. di W-K ha la $\sum_{-\infty}^{+\infty}$, non ha una finestra sopra ed è proprio quella ^{finestra} che fa sì che se calcolo la funz. di autocorr su una finestra limitata, quando ne faccio la transf. di Fourier rimane una funzione pari (e quindi la transf. reale), ma non positiva.

Questo si nota nelle slide 11 e slide 13 \rightarrow le aree tra H e p, nelle slide 13, sono aree che originariamente ~~avrebbero~~ avuto valori negativi ma non bisogna fare correlogrammi con valori negativi perché \nexists la potenza negativa. Allora:

- o come nel grafico nelle slide 13, li ribalto i valori negativi e li faccio diventare positivi (e quindi prendo il modulo del correlogramma);
 - oppure i valori negativi li metto a zero \rightarrow tranco la parte negativa.
- Sono 2 scelte valide entrambe, ma la potenza negativa NO!

Questo è uno degli effetti del calcolo numerico fatto al finito, su una lunghezza limitata.

2) ~~Si~~ ~~metodo~~ ~~per~~ ~~ritardi~~ ~~negativi~~ $(-(N-1) \leq m < 0)$:

$$\hat{E}_{x,x}[m] = \frac{1}{N-|m|} \sum_{n=0}^{N-|m|-1} x^*[m+1+n] x[n]$$

Nella pratica, una volta calcolata per ritardi positivi, il calcolo per ritardi negativi è molto più semplice.

Durante il ~~laboratorio~~ 3 dobbiamo fare una funzione Matlab che, dato un ~~sgn~~, ~~me~~ ~~calcola~~ la funzione di autocorrelazione \Rightarrow bisogna trovare una formula esplicita in funzione di ~~reale~~.

Matlab ha già una funzione che fa questo e si chiama: **XCORR** e serve per ~~plottare~~, ~~generalmente~~, una funzione di ~~cross-correlazione~~ (ma se do in ingresso lo stesso ~~sgn~~, ~~calcola~~ la funzione di autocorrelazione) = ~~XCORR~~ sarà la verifica verso cui volterò se la funzione è corretta o

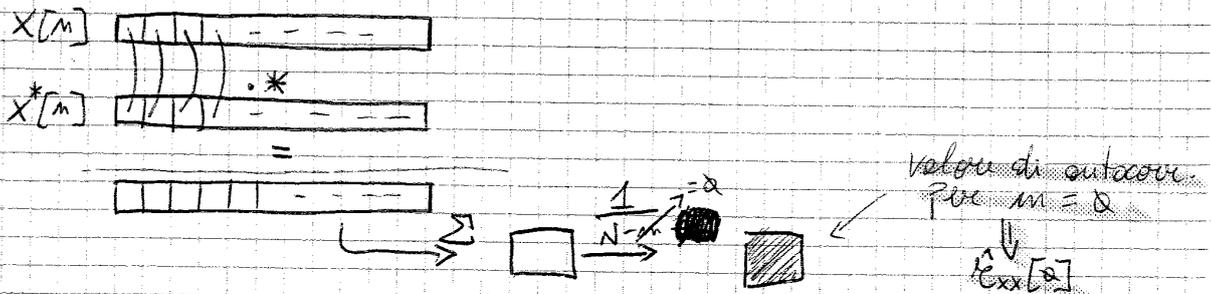
Considerazioni: abbiamo un ~~sgn~~ e supponiamo che sia la nostra sequenza numerica $x[n]$ fatta da N campioni e devo calcolare la seq. di autocorrelazione



Allora: dato un ritardo discreto m (positivo), \forall valore di m dobbiamo andare e calcolare il valore della funz. di autocorr. secondo la formula 1).

Il ritardo minimo, positivo, che posso imporre è 0. Allora:

- $m=0$ \rightarrow la formula dice che devo prendere il segnale e questo segnale lo devo ritardare di un numero m di campioni. Ma se $m=0$ non lo ritardo, e lo moltiplico per il suo complesso coniugato, allora



NOTA

Se Matlab:

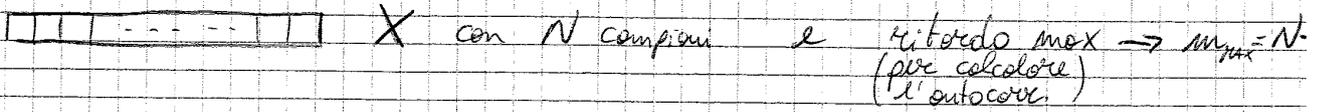
- $a * b \rightarrow$ se uno è riga e l'altro è un vett. colonna \Rightarrow ottengo una scala
- $a . * b \rightarrow$ moltiplico elemento per elemento e ottengo un vettore.

anche se i spn biologici sono rasti, tanto non costa niente di più.

L'istruzione Matlab per calcolare il complesso esposto di un vettore è:

conj.

Supponiamo di avere:

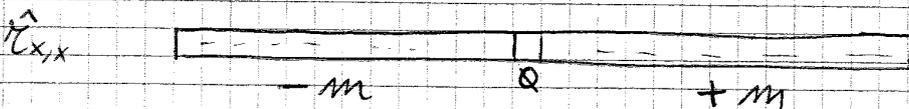


Se calcolo le funz. di autocorr. per tutti i ritardi e dispersione (da α a $N-1$), il vettore delle funz. di autocorr. sarà grande $2N-1$ (N reali positivi, N rit. negativi, e lo α) \rightarrow questo vale sempre, cioè: se

supponiamo di calcolare le funz. di autocorr. per un ritardo $m' \leq N-1$, vuol dire che le funz. di autocorr. è un vettore formato da un numero pari e $(2m'-1)$ campioni. Questo vuol dire che se ho N , non devo calcolare l'autocorr. per tutti i ritardi, ~~il ritardo max~~ il ritardo max per cui calcolo l'autocorr. lo decido io e in stime spettro questo è un parametro cruciale \Rightarrow un altro motivo per cui è imp. mostrare bene la seq. di autocorr. è che manualmente prendo il spn che è formato da un numero molto grande di campioni, calcolo la sequenza di autocorr. su tutti i ritardi possibili, ne faccio le trasform. di Fourier e ottengo una roba inutile (quasi sempre).

Nella pratica, molto spesso, conviene calcolare le funz. di autocorr. per un ritardo massimo minore del numero di campioni.

Capito che, se questo è:



il vettore di \hat{r}_{xx} , lo α è il centro e da un lato ci sono tutti i ritardi positivi e da un lato tutti quelli negativi.

Questo stimatore $\hat{r}_{xx}[m]$ delle seq. di autocorr. non è unico ma ne è un altro ALTERNATIVO (Slide 34), fatto così:

$$\hat{r}_{xx}[m] = \frac{1}{N} \sum_{n=0}^{N-m-1} x[n+m] x^*[n] \quad , \text{ per: } 0 \leq m \leq N-1$$

$$\hat{r}_{xx}[m] = \frac{1}{N} \sum_{n=\alpha}^{N-|m|-1} x^*[n+|m|] x[n] \quad , \text{ per: } -N \leq m < \alpha$$

CAMBIA...

me faccio le stime con uno stimatore che scelgo io su un sottocampione, si dice che questo stimatore $\hat{\theta}$ è non polarizzato, quando: $E\{\hat{\theta}\} = \theta$

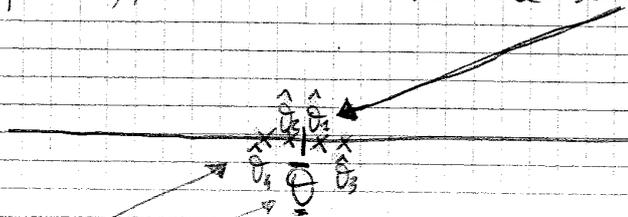
A noi piace avere uno stimatore NON POLARIZZATO \Rightarrow la polarizzazione è certamente uno dei primi parametri con cui si sceglie uno stimatore.

Io so già che lavorerò al finito e voglio che lo stimatore, al finito mi dia delle prestazioni identiche a quello che mi dà lo stimatore all'infinito.

\Rightarrow Se il valore ottenuto di $\hat{\theta}$ non coincide con la quantità da stimare, si dice che lo stimatore è POLARIZZATO, cioè la differenza tra la quantità reale e il valore ottenuto dello stimatore si chiama polarizzazione (o errore di stima).

• Altro parametro è la CONSISTENZA dello stimatore.

Supponiamo che la quantità da stimare (cioè il valore medio su tutta la popolaz. definito prima), su una scala lineare è qui come valore numerico



e supponiamo di avere uno stimatore che su diversi campioni, mi restituisce questi valori.

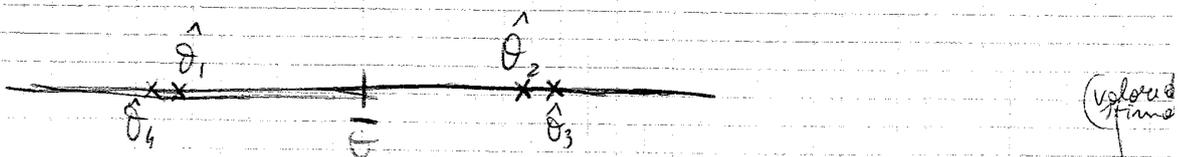
\Rightarrow se calcolo il valore medio di $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ e $\hat{\theta}_n$ ottengo $\bar{\theta}$ e quindi questo stimatore è ovviamente non polarizzato.

Ma è uno stimatore CONSISTENTE? consistente vuol dire che su campioni diversi, mi ha restituito sempre lo stesso stima? No!

Lo stimatore non polarizzato e perfettamente consistente, è quello stimatore che su ogni campione mi restituisce sempre esattamente $\theta \rightarrow$ NON \exists !!

Quindi ci accontentiamo del fatto che questo stimatore restituisca dei numeri che non sono sempre gli stessi, ma che sono centrati e vicini al valore che devo stimare.

Supponiamo un altro stimatore:

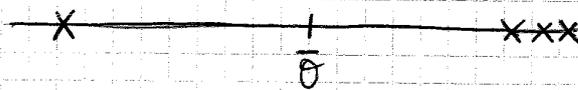


anche in questo caso, il valore ottenuto di questo stimatore coincide con θ

Tra i due esempi, è meglio il 1° stimatore perché è vero che tutti e due sono non polarizzati, ma il 1° ha dei valori molto più vicini al valore da stimare e quindi la varianza della stima misura la consistenza dello stimatore

⇒ più basso è la varianza dello stima, più consistente è lo stimatore. L'ideale sarebbe varianza = 0, ma non si può ottenere.

I stimatori non polarizzati e consistenti; I stimatori non polarizzati ma non consistenti ed I stimatori polarizzati e consistenti ma la maggior parte sono polarizzati e non consistenti → caso classico:



Riepilogo:

- Per verificare se uno stimatore è polarizzato, calcolo il VALORE ATTESO
 - Per verificare quanto è consistente uno stimatore, calcolo la VARIANZA
- oss. anche se statisticamente la consistenza è definita come: $P\{|\hat{\theta} - \theta| \leq \epsilon\} = 1 - \alpha$, se la probabilità che lo stima che faccia disto dalle quantità che devo stimare di un valore piccolo o piccolo, siamo nel caso di stimatore consistente.

Allo (S. 33) Allora, proviamo a calcolare valore medio ^(atteso) e varianza per capire se lo stimatore $\hat{x}_{x,x}$ è polarizzato e/o meno consistente.

Per gli stimatori statistici, andare a calcolare la polarizzazione è facile (perché valore atteso è un'operazione lineare), mentre calcolare la varianza non è facile (il più delle volte ci si basa solo se si fanno hp su quello che c'è all'interno delle Σ , in questo caso di quello che è il μ) ⇒ calcolare il valore atteso di $\hat{x}_{x,x}$ è facile, mentre la varianza la scriviamo in forme chiuse solo supponendo che x sia una variabile casuale, gaussiana, ecc...

Però ci dà un'indicazione di massima delle prestazioni di questo stimatore. Le prestazioni quali sono?

• CALCOLO VALORE ATTESO:

$$E\{\hat{x}_{x,x}[m]\} = \frac{1}{N-m} \sum_{n=0}^{N-m-1} E\{x[n+m] x^*[n]\} = r_{x,x}[m]$$

⇒ ho uno stimatore, il cui valore atteso coincide con la variabile da

Invece, se vado a calcolare la varianza, succede che ho:

$$\text{var} \{ \hat{r}_{x,x}^v[m] \} \approx \frac{1}{N} \sum_{k=-\infty}^{+\infty} (r_{x,x}^v[k] + r_{x,x}^v[k+m] r_{x,x}^v[k-m])$$

⇒ è ancora vero che all'aumentare del ritardo, aumenta la varianza però a differenza della varianza precedente, succede che all'aumentare del n° di campioni (N), la varianza va a 0 ⇒ è polarizzato (perché è asintot. non polarizzato), ma è asintoticamente consistente, nel senso che ha una varianza nulla per $N \rightarrow \infty$ (cosa che lo stimatore $\hat{r}_{x,x}$ non aveva).

⇒ questo stimatore $\hat{r}_{x,x}^v$ ha delle prestazioni peggiori rispetto a quello che stima, ma ha una variabilità più bassa. Quindi di volte in volte sceglieremo lo stimatore che ci fa più comodo → se ho un problema di stima spettrale dove so già che la variabilità della stima è un problema, devo scegliere a mente uno stimatore a bassa variabilità (cioè quello consistente); se invece so che ad es. può essere un problema per la stima dell'autocorrelazione, la polarizzazione, perché voglio valori precisi ⇒ scelgo quello non polarizzato.

Tutto non si può avere...
(Sudoe 36)

C'è ancora un altro problema per cui è stato introdotto questo stimatore alternativo → il problema è che dalla teoria si sa che la funz. di autocorrelazione ha un massimo assoluto nell'origine. ⇒ si è visto che per alcune sequenze numeriche (un po' strane) questo non era corretto.
ES: prendo una seq. numerica fatta da 3 numeri:

$$1 ; 1, 1 ; 1$$

Se calcolo i valori della seq. di autocorrelazione per ritardo 0, 1 e 2 con lo stimatore $\hat{r}_{x,x}$ (NON polarizzato), ottengo:

$$\hat{r}_{x,x}[0] = 1, 07 ; \hat{r}_{x,x}[1] = 1, 1 ; \hat{r}_{x,x}[2] = 1, 0$$

Il massimo ce l'ho per ritardo 1, non in 0 ⇒ questo capita per sequenze particolari (anche corte), che lo stimatore desse dei risultati non fisicamente accettabili. Questo non succede con lo stimatore polarizzato $\hat{r}_{x,x}^v[m]$, perché tanto divido sempre per N e quindi i ritardi successivi devono essere per forze più bassi.

25/3/15

(SLIDE 38) Una volta stimata la ACS, il correlogramma lo ricavo seguendo il teorema di W-K \rightarrow abbiamo detto che questo teorema ha una doppia valenza: ci definisce lo spettro, ma possiamo anche leggerlo come criterio operativo di stima. Allora dato che il th di W-K dice:

$$P_{x,x}(f) = T \sum_{m=-\infty}^{+\infty} r_{x,x}[m] e^{-j2\pi f m T}$$

cioè che la PSD del segnale, è la trasform. di Fourier delle funz. di autocorr. \Rightarrow dato che noi non abbiamo la funz. di autocorr. tra $-\infty$ e $+\infty$ e l'unica cosa che si può fare per avere una stima della PSD è sostituire alla $r_{x,x}$, la stima fatta (a partire dal spu): $\hat{r}_{x,x}$. Allora:

$$\hat{P}_{x,x}(f) = T \sum_{m=-L}^{+L} \hat{r}_{x,x}[m] e^{-j2\pi f m T}$$

Questo è il CORRELOGRAMMA \Rightarrow la trasform. di Fourier delle stime delle funz. di autocorr. è una stima della PSD del spu e lo chiamiamo correlogramma.

La differenza è che $\hat{r}_{x,x}$ non è la funz. di autocorr. reale (perché noi dobbiamo lavorare al finito e quindi lo Σ va da - un rit. max a + un rit. max).

NOTA: Generalmente il rit. max è \leq ed $\frac{1}{10}$ del numero di campioni (N) di cui è costituito il spu (per ridurre la varianza dell'autocorr.): questo fattore $\frac{1}{10}$ lo vedremo più avanti (FATTORE DI QUALITÀ della stima), ma dice una cosa semplice, cioè che se voglio avere una misura numerica di quanto una stima è affidabile o meno, bisogna fare i conti con la sua variabilità e con la quantità da stimare. Se ad es. si ha una variabilità di stima di 20 cm quando devo misurare la lunghezza di un tavolo, è tantissimo; ma se ho una variabilità di 20 cm quando devo misurare la lunghezza dell'autoritratto, è accettabile \Rightarrow vedremo più avanti che si introduce il fattore di qualità che rapporta la variabilità della stima alla quantità da stimare. Dalla teoria che c'è dietro il fattore di qualità, c'è stata una proposta \rightarrow un buon compromesso per tenere a bada la variabilità di stima nel caso del correlogramma, è usare un n° di ritardi non troppo alti su cui calcola la sequenza di autocorrelazione; per molti segnali si bene stare sotto $\frac{1}{10}$ della lunghezza del numero di campioni di cui è costituito il segnale. Però non è una regola ferrea, ma se

(Slide 40)

Se invece di calcolare il covariogramma a partire dallo stimatore non polarizzato della ACS, calcolo il covariogr. associato allo stimatore polarizzato, non cambia tanto perché la definizione è la stessa e il covariogramma è:

$$\hat{P}_{xx}^v(f) = T \sum_{m=-L}^{+L} \hat{r}_{x,x}^v[m] e^{-j2\pi f m T}$$

Se calcolo la sua polarizzazione \Rightarrow ne faccio il valore atteso, ricordando che lo stimatore polarizzato è uguale a quello non polarizzato tramite un fattore di conversione \Rightarrow quando calcolo il valore atteso di $\hat{P}_{xx}^v(f)$, ottengo:

$$E\{\hat{P}_{xx}^v(f)\} = T \sum_{m=-L}^{+L} \left(1 - \frac{|m|}{N}\right) r_{x,x}[m] e^{-j2\pi f m T} = P_{xx}(f) * \frac{2}{L} D_c^2\left(\frac{f}{2}\right)$$

\Rightarrow quando sposto infatti il valore atteso all'interno della sommatoria, bisogna calcolare il valore atteso dello stimatore polarizzato, che è uguale a:

$\left(1 - \frac{|m|}{N}\right) \cdot r_{x,x}[m]$. La \sum è sempre finita (tra $-L$ ed L) \Rightarrow succede che se vedo e vedo la differenza tra il caso ^{non} polarizzato (Slide 39) e questo polarizzato, se disegno la funzione di questo caso polarizzato in funzione del ritardo L , ottengo proprio una finestra triangolare \rightarrow ~~1 - |t|~~ $1 - |t|$ è l'espressione della finestra triangolare, che è detta finestra di Bartlett che funziona un po' meglio delle finestre rettangolare perché ha i lobi secondari che sono un po' più bassi.

Allora, quando facciamo la transf. di Fourier abbiamo la convoluzione tra lo spettro del segnale e lo spettro della finestra triangolare.

\Rightarrow Per quanto riguarda la polarizzazione, è meglio questo caso, è meglio cioè la finestra triangolare perché i lobi secondari sono più bassi e chi è responsabile del leakage di potenza sono i lobi secondari delle finestre.

Allora, se uso la versione polarizzata dello stimatore, ottengo comunque un covariogramma polarizzato, ma meno polarizzato del caso in cui utilizzassi la versione non polarizzato dello stimatore \rightarrow questo è il motivo per cui si preferisce spesso calcolare il covariogramma a partire dalla versione polarizzato delle sequenze di autocorrelazione. Infatti, il covariogramma è sempre polarizzato, in più

lo spettro della finestra (cioè la transf. di Fourier della finestra che ho considerato).

In generale allora questo covelogramma è polarizzato, ma se scelgo una finestra dove i lobi secondari dopo un po' vanno a zero e non sono di valore troppo elevato e calcolo il limite per $L \rightarrow \infty$ (cioè per ACS infinitamente lunga, che vuol dire anche σ_{sp} infinitamente lungo), scopro che questo valore ottiene diventa uguale allo spettro reale del segnale \Rightarrow in generale si può dire che questo stimatore è asintoticamente non polarizzato. Cioè se faccio il $\lim_{L \rightarrow \infty}$, otteniamo il valore corretto, ovvero otteniamo lo spettro del $\sigma_{\text{sp}} \rightarrow$ più di così non si ottiene. Allora se ho problemi di polarizzazione, l'unica cosa che si può fare è prendere lo stimatore della ACS, metterci sopra una finestra ed effettuare la transf. di Fourier \Rightarrow questo mi garantisce una polarizzazione minore, ma lo stimatore rimane polarizzato con il covelogramma, meglio di così non si fa.

Per questo riguarda invece la variabilità (cioè la consistenza), come si è già visto questa è legata al n° di ritardi e si è detto che maggiore è L , maggiore è la variabilità \Rightarrow ottengo un covelope troppo variabile, riduco il n° di ritardi su cui ho calcolato la ACS

~~Operativamente non esiste una funzione Matlab che mi aiuti, ma tutto non serve perché il covelogramma è semplicemente la transf. di Fourier della ACS \Rightarrow casomai esiste una routine che calcola la ACS.~~

Quindi, una volta che ho lo stimatore polarizzato o non polarizzato della ACS, decido il n° di ritardi, decido se finestrare o meno e poi ne faccio la transf. di Fourier \Rightarrow ~~NON SERVE NESSUNA ROUTINE PARTICOLARE~~

oss.: in genere, per fare una stima spettrale su un σ_{sp} reale, il metodo del covelogramma non lo è molto apprezzato \rightarrow al compito è usato da tutti il periodogramma. Un motivo può essere quello per cui il covelogramma può avere valori negativi, invece il periodogramma è $|f|^2$ ed è reale, positivo. Però il covelogramma concettualmente è più semplice, ci sono molte meno cose da sapere rispetto al periodogramma: solo il n° di ritardi e la finestra per tenerne e badare, rispettivamente la variabilità e la polarizzazione.

Se calcolo la varianza, trovo che questa varianza non va mai a zero nemmeno asintoticamente \Rightarrow è uno stimatore poco consistente, perché la varianza è molto elevata.

La varianza non va mai a zero ~~perché~~ perché anche se $N \rightarrow \infty$, il termine nelle parentesi quadre non si annulla mai e quindi ha sempre una varianza che è dell'ordine del quadrato di quello che si deve stimare.

$$\text{Var} \left\{ \hat{P}_{xx}(f) \right\} = P_{xx}^2(f) \left[1 + \left(\frac{1}{N} \frac{\sin(2\pi fT)}{\sin(\pi fT)} \right)^2 \right]$$

\Rightarrow il periodogramma semplice ha delle prestazioni che abbiamo già visto, perché lavorando al finito abbiamo molti casi.

Ancora oggi, la stima spettrale, è difficile da fare perché non tiene conto del fatto che sto lavorando con robe che ha prestazioni limitate: bisogna fidarsi, ma non tanto; oppure bisogna scegliere stimatori di cui ci si può fidare un po' di più. Allora bisogna cercare di abbattere la variabilità di stima: nel periodogramma, e differenza del cosinogramma, non si ha come ACS (e quindi non si può ridurre il numero di ritardi) e quindi per abbattere la variabilità di stima, come si fa? Il periodogramma semplice si chiama così perché non fa niente né per la polarizzazione, né per la variabilità; ma nel corso degli anni sono stati proposti un certo numero di periodogrammi alternativi, con il compito proprio di migliorare le caratteristiche non proprio ideali dello stimatore basato sul periodogramma semplice.

In particolare il punto focale su cui tutti si sono concentrati, è la variabilità, cioè la consistenza della stima \rightarrow è stato fatto un grande sforzo per rendere gli stimatori più consistenti. Per la polarizzazione tanto, se lo voglio migliorare, devo finire il segnale (mentre per il cosinogramma finisco la ACS), prima di fare la trasform. di ff.
 \Rightarrow la polarizzazione non ci dà molti problemi, basterebbe scegliere la finestra più adatta. Il vero problema è la variabilità della stima che rende quello che ottengo, inaffidabile.

(SioE 53) \rightarrow Questi metodi sono stati proposti non per migliorare la polariz., ma per ridurre la variabilità: METODO di DANIELL, METODO di BARILETTI, METODO di WELCH \Rightarrow il vero motivo che c'è dietro, è quello di ottenere

⇒ se ho un segmento da 256 campioni (~~ovvero~~ o meno non ci importa), campionato a 64 Hz, la risoluz. spettrale teorica è: 250 mHz, perché

$$T = \frac{256}{64} = 4 \text{ secondi} \Rightarrow \Delta f = \frac{1}{T} = 250 \text{ mHz}$$

Se non voglio WELCH, ma voglio BARTLETT, cioè:

$f_c = 64 \text{ Hz}$
 $\Delta f = 0,5 \text{ Hz}$ ⇒ devo usare: $W = \text{boxcar} (128)$
 $\theta = \emptyset$ # pochi non c'è sovrapposizione

Se ancora voglio il periodogramma di SEKUSTER (o semplice):

$W = \text{boxcar} (\text{length}(x)) \rightarrow$ perché ci devo mettere solo la finestra rettangolare, lunga quanto il segnale
 $\theta = \emptyset$

Ovviamente è possibile mettere una finestra boxcar di una certa lunghezza e mettere anche overlap, ma non si chiama né Bartlett, né Welch, né semplice → si chiama solo periodogramma.

L'istruzione pwelch ha esattamente gli stessi parametri di psd, ma cambia l'ordine in cui si danno, ma la funzione è la stessa (cambia solo le intesi).

5.4.7) RISOLUZIONE SPETTRALE TEORICA ED APPARENTE

Si è detto che come regole generali bisogna distinguere tra risoluzione spettrale teorica (che dipende unicamente da T , che è la durata temporale del pezzo di cui facciamo lo trasf. di Fourier) e risoluzione spettrale apparente (che dipende solo dal no di punti che vogliamo rappresentare sull'asse delle frequenze e che è governato dal parametro NFFT).

oss:

NFFT è un numero che teoricamente può assumere qualsiasi valore, ma dato che in un algoritmo di calcolo bisogna risolvere un sistema lineare di un certo ordine dato da NFFT, alcune routine (tra cui psd) di ~~Matlab~~ pongono dei vincoli, ad es: si può prendere NFFT più grande (a piacere) della lunghezza dei campioni del segnale, ma non si può prendere più piccolo (perché genererebbe un errore).

Il problema è che teoricamente all'interno di NFFT ci possiamo mettere quello che vogliamo, ma in pratica dobbiamo rispettare un certo numero di regole ⇒ succede che quasi tutte le routine non consentono di avere una risoluzione spettrale apparente inferiore a quella teorica.

Se ho un pezzo di segnale che mi garantisce una risoluz. spettrale teorica di 1 Hz, vuol dire che sull'asse delle frequenze i punti originali sono equidistanti tra di loro di 1 Hz. ⇒ pochi doveri rappresentarli con

alle fine del segnale, cioè da $X[N]$ a $X[2N-1]$ e poi calcoliamo la trasformata di Fourier discreta di questa sequenza (DFT):

$$X[k] = \sum_{m=0}^{2N-1} X[m] e^{-j \frac{2\pi mk}{N}}$$

La sommatoria va da 0 a $2N-1$, cioè è una sommatoria fatta del doppio di punti, perché ho aggiunto N zeri \Rightarrow da origine a un sistema che ha una dimensione doppia di quella di prima e quindi non ho più N equazioni ed N incognite, ma ho $2N$ equazioni e $2N$ incognite (ho raddoppiato il n° di punti). Però alcune di queste equazioni sono 0, perché $X[m]$, per tutti i valori che vanno da $m=N-1$ in avanti, è nullo \Rightarrow (siccome) se effettuo un cambio di variabile, ponendo $K=2k$, scopro che quel sistema sta restituendo tutti i valori (cioè tutti i campioni) di posto pari, e quelli di posto dispari non sono altro che campioni interpolati da una funzione $\frac{\sin x}{x}$. Questa funzione $\frac{\sin x}{x}$ è quella che abbiamo

chiamato kernel di Dirichlet, cioè lo spettro della finestra boxcar \Rightarrow in realtà posso ottenere tutti i campioni che voglio, ma i campioni originali rimangono, mentre quelli che sono aggiunti in più tramite FFT, sono solo interpolati tramite un interpolante di tipo $\frac{\sin x}{x}$ \rightarrow è per questo che si chiama risoluzione spettrale opposta, perché sembra che i campioni in frequenze siano più vicini tra di loro (e quindi ci sia più risoluzione), ma è finto, perché ho solo interpolato i valori originali.

(Siccome)

Questo esempio fa vedere abbastanza bene questa cosa:

a) \rightarrow questo è il modulo della trasformata di Fourier di un segnale a 16 punti, che contiene 3 sinusoidi (è lo spettro teorico "continuo");

b) \rightarrow questo è la trasformata di Fourier discreta di questo segnale, senza ZERO PADDING, ma usando solo i 16 punti che avevo nel dominio del tempo. Quanti punti ci sono in frequenza? $9 \rightarrow$ il primo campione

che la risoluzione spettrale teorica è l'inverso della lunghezza del pezzo che trasformo con Fourier \rightarrow nel caso del periodogramma, è il sgu; nel caso del coelogramma, è lo ACS. Quindi, procedendo sui retardi, determino la risoluz. spettrale teorica. Mentre quella apparente? NFFT esiste perché il periodogramma applica una trasform. di Fourier, come anche il coelogramma; ma il numero di punti, è quello che voglio, in frequenza e lo determino proprio la trasf. di Fourier che diventa una griglia continua. \Rightarrow quando calcolo la FFT dello ACS, lì devo definire il numero di punti giusto per avere quella risoluzione spettrale apparente che voglio imporre o che viene richiesta.

\Rightarrow RISOLUZIONE SPETTRALE TEORICA \rightarrow lunghezza del pezzo di cui si fa FFT
RISOLUZIONE SPETTRALE APPARENTE \rightarrow NFFT
UGUALE INDIPENDENTEMENTE DA CORRELOGRAMMA O PERIODOGRAMMA

Slide 42) PRODOTTO TEMPO - LARGHEZZA DI BANDA

Abbiamo detto che:

$$\Delta f = \frac{1}{T}$$

\uparrow risoluz. spettr. teorica \uparrow lunghezza temporale del segnale

Questa regola (che deriva dalle proprietà delle trasf. di Fourier) è una simplificazione di una regola empirica. ~~...~~ In realtà se vado a vedere cosa dice la teoria delle trasform. di Fourier reali, non è proprio vero che la risoluz. spettrale teorica è proprio l'inverso della durata temporale, perché ciò è vero considerando quello che viene chiamato SUPPORTO EQUIVALENTE e BANDA EQUIVALENTE.

T_e

B_e

Slide 43)

Supponiamo di avere un segnale $x(t)$ (grafico (a)) $\Rightarrow T_e$ è il supporto temporale del segnale quando lo "impacchetto" e diventa quadro, cioè può diventare anche ~~...~~ rettangolare, ma diciamo che mantengo fisso la ...

se ho una finestra che ha un lobo principale molto largo, questo spettro viene convoluto per tutte le S di Dirac \Rightarrow vuol dire che il lobo principale, è sempre centrato dove c'è la frequenza corretta (cioè sulla S di Dirac), ma non è più stretta, non ci più tutta la potenza intorno a quella frequenza; la potenza è spalmata su un intervallo di freq. che è più grande e quindi non si può dire di avere la stessa risoluzione in frequenza, perché per me tutte le frequenze che cadono all'interno del lobo principale, sono tutte insieme (le vedo come una frequenza). Allora quella larghezza diventa la misura di quella che effettivamente è la mia banda equivalente (quando ho una finestra di stima spettrale).

(Si: 05 15)

Allora, per spu canonici, la relazione $T_e \cdot B_s = 1$, si trasforma in:

$$\boxed{Q \cdot T_e \cdot B_s = 1}$$

Questo è drammatico dal punto di vista delle prestazioni, perché dato che si è detto che la stima è di buona qualità quando il valore di Q è piccolo (< 1) \Rightarrow vuol dire che se ad es. $Q = 0,1$, $T_e \cdot B_s = 10$. Quindi, se faccio con 1 secondo di spu ovvero una $\Delta f = 1 \text{ Hz}$, edesso a 1 secondo di spu ho la risoluzione spettrale di 10 Hz. Allora questo regale sta dicendo che se prima potero considerare punti equivalenti 1 Hz (come i veri punti (quelli che effettivamente mi danno la risoluz. spettrale teorica), edesso per non incorrere in fluttuazioni canonici, se $Q = 0,1$ io considero significativi (e quindi come punti che mi danno la risoluzione del spu), i punti che distano tra di loro 10 Hz. Allora se volessi di nuovo avere una risoluz. spettrale teorica di 1 Hz, mi servono 10 secondi di spu \rightarrow questo è drammatico perché solo così posso dire che ho una stima su un spu canonico di buona qualità. Con 1 secondo di spu non posso trovare 1 Hz di risoluz. spettrale, ~~ma~~ ma solo nelle migliori delle hp ignorando Q , in nella pratica non è così perché se Q lo ignoro, mi ritrovo dell

fluttuazioni casuali che interpretato come risoluzione vera dello spettro è fatto molte volte è sbagliato.

Altra ~~nella pratica~~ ^(conoscendo i conti) ~~quando facciamo i conti, i prezzi ignorare Q, ci facciamo finta che: $Q=1$~~ \Rightarrow si persiste la relazione che $T \propto B^2$ che se effettivamente vogliamo essere sicuri di tenere conto delle qualità della firma, Q bisogna considerarlo del valore corretto \Rightarrow non si può più distinguere un grafico con tutti gli evidenti, che se fatto un grafico in cui i punti sono equidistanti (come nell'es. di prima, ed es., di 50 Hz). tutti gli altri punti non servono a niente, dobbiamo assumere che siano tutti punti casuali.

Si vedrà che nella pratica, ci sono casi in cui ignorare Q non comporta nessun problema; altri casi in cui, invece, ignorare il valore di Q cambia decisamente le cose

SCHEMA RIASSUNTIVO \rightarrow 2 CASI :

1) **IGNORO Q** (per decisioni mentre o su indicazioni) $\Rightarrow Q=1$

Vol dire che lo stima è ad alta variabilità (SICRA non CONSISTENTE)

2) **CONSIDERO Q**
 $\left\{ \begin{array}{l} - se fissa Q e B e ottiene T \\ - se fissa B e T e ottiene Q \end{array} \right.$
 ~~fora Q e B e T~~

\rightarrow dipende da cosa devo fare o cosa chiedo il testo

Si hanno 3 parametri: Q, T, B \rightarrow 2 vanno imposti e 1 va di conseguenza.

\Rightarrow Se ad es. mi viene dato: $Q = 10^{-3}$, effettua firma spettro partendo Δf di almeno 1 Hz, come si fa??

$$T = \frac{1}{\Delta f \cdot Q} = \frac{1}{10^{-3}} = 1000 \text{ secondi} \Rightarrow \text{con queste condizioni}$$

devo prendere un riferimento temporale lungo 1000 secondi, fare periodo e correlazione e garantire che sul piano di prova disegno lo spettro, ogni punto

fluttuazioni casuali che interpretato come risoluzione vera dello spettro e
questo molte volte è sbagliato.

Allora nella pratica ^(esistono 2 casi) quando facciamo i conti, i prezzi ignorare Q , cioè
facciamo finta che: $Q=1 \Rightarrow$ ti esprimiamo la relazione che $T \cdot B_s = 1$
Ma se effettivamente vogliamo essere bravi di tenere conto della qualità
della stima, Q bisogna considerarlo del valore corretto \Rightarrow non si può
più disporre un grafico con tutti gli andamenti, come se fatto un grafico in
cui i punti sono equidistanti (come nell'es. di prima, ad es., di 50 Hz) e
tutti gli altri punti non servono e niente, dobbiamo assumere che siano
tutti punti casuali.

Si vede che nella pratica, ci sono casi in cui ignorare Q non comporta
nessun problema; altri casi in cui, invece, ignorare il valore di Q cambia
decisamente le cose

SCHEMA RIASSUNTIVO \rightarrow 2 CASI :

1) IGNORO Q (per decisioni morali o su indicazioni) $\Rightarrow Q=1$

vol dire che la stima è ad alta variabilità (stima non
CONSISTENTE)

2) CONSIDERO Q \Rightarrow dipende da cosa devo fare o cosa chiede il testo

Si hanno 3 ^{Altre questo potrebbe essere:} ~~fora~~ ^{fora stima spettro non promettuto,} \Rightarrow 2 vanno imparti e 1
vo di conseguenza ^{dato un Δf , potremmo un}
fattore di qualità: $Q = 0.5$.

\Rightarrow Se ad es. mi v_i ^{Come faccio? $T \cdot B_s$ deve essere = 10, effettuare stima spettro}
potremmo Δf ^{\Rightarrow ad es. prendo 50 sec. di Δf e} T ^{ottenso una risoluz. spettro di 200} z ^{si fa??}

$T = \frac{1}{\Delta f \cdot Q} =$ ~~anni~~ ^{anni}, in questa ~~area~~ ^{area} \Rightarrow con queste condizioni

altro prendere un ^{1 secondo, se considero Q ,} 1000 secondi, ^{dipende cosa devo fare.}
o envelope, ^{- può vedere se finisce Q a T} 1000 secondi, ^{e ottenere B}
e garantire che sul piano di ^{ogni} 1000 secondi, ^{fora prevedere}
ogni punto ^{lo spettro, ogni punto}

Ci piace tanto: può essere negativo nella pratica \Rightarrow se calcolo tanti correlogrammi su pezzi di z_{pu} e poi faccio la media, se mi capita casualmente l'intervallo negativo, faccio la media con una cosa negativa e il risultato è ~ 0 . Quindi nella pratica, il fatto di lavorare al finito, e per le cattive prestazioni del correlogramma, fa sì che la media di correlogrammi è un'operazione che è meglio non fare.

Inoltre non ci serve nemmeno la media, perché Blackman e Tukey hanno detto che la variabilità si può tranquillamente ridurre, riducendo il no di ritardi, mentre la polarizzazione si riduce filtrando la sequenza di autocorrelazione.

Quindi: gli approcci tra correlogrammi e periodogrammi sono tecnicamente equivalenti, nella pratica c'è chi trova più facile fare una cosa, chi un'altra, ma è solo una questione di prendere la mano.

Ora andiamo a vedere come applicare nella pratica tutta questa teoria introducendo caratteristiche dei z_{pu} bianchi \Rightarrow le caratteristiche dei z_{pu} non dovrebbero essere nuove, ma bisogna capire come fare le cose nella pratica ed è che si sa che non è tutto così semplice (non posso dire che prendo il z_{pu} e ne ricavo lo spettro, ne ricavo uno spettro brutto, ovvero una stima con tutte le difficoltà del caso). Allora è secondo di quello che ci devo fare con il z_{pu} , devo capire quali ~~sono~~ ^{le} caratteristiche della stimolazione che mi interessano di più.

Lo fanno inclinare in ottimi diversi (ciò sono a cavallo tra diversi neuroni).
Sempre nel disegno allo slide 5, in alto si vedono la ~~superficie~~ superficie
esterna del cranio $F \rightarrow$ della corteccia cerebrale all'osso occipite, c'è un
certo numero di membrane (p.a. madre, ~~altro madre~~, meningi, ecc...) \Rightarrow quelle
è la parte più esterna. La parte in basso dove c'è la V, sarebbe il
ventricolo cerebrale: i ventricoli sono strutture cave all'interno del
cervello che contengono liquido cerebrospinale e hanno compiti diversissimi.
• Il compito più importante è che, essendo pieni di liquido ed essendo
che il liquido è incompressibile, determinano la pressione intracranica,
grazie alla quale il cervello non si muove, sta fermo nella teca
occipite e non collasa.

• Il secondo compito altrettanto importante, è che il liquido cerebrospinale
è un liquido molto particolare che è fatto apposta per scambiare sostanze
con la materia cerebrale. E quindi sono raccolte di liquido con
funzioni di tipo meccanico e funzioni di tipo metabolico (cioè
afferente che efferente \rightarrow cioè portano sostanze e portano via in parte
anche sostanze di scarto).

Da questa figura si vede che quelli che noi chiamiamo neuroni, cioè la parte
nobile delle cellule cerebrali, non è mai direttamente a contatto né con
le membrane esterne del cervello, né con il ventricolo \rightarrow i neuroni stanno
all'interno per i fatti loro e quindi né toccando la parte esterna né
la parte interna, come fanno a scambiare sostanze e quant'altro? Non
lo fanno direttamente loro, ma altri neuroni di supporto in verde,
che si chiamano astrociti.

La barriera emato-encefalica (BBB: Blood-Brain barrier \rightarrow barriera
sangue - cervello) fa sì che tutto quello che va nel circolo ematico, non
~~arriva~~ arriva ai neuroni \rightarrow c'è questa barriera di mezzo che fa sì che
molte sostanze in circolo, non vedono assolutamente ai neuroni: questa è una
condizione di protezione del sistema cerebrale. Per cui i neuroni più

Poi ci sono altre patologie dove muoiono gli oligodendrociti e quindi gli assoni diventano nudi, non hanno più le guaine mieliniche → la sclerosi multiple, è un grande problema perché in Italia come nel resto del mondo, è una malattia giovanile (insorgenza media intorno ai 25 anni) ⇒ progressivamente queste guaine mieliniche vanno deteriorate dai macrofagi, cioè il sistema immunitario si mangia le guaine mieliniche; ~~Ma~~ perché ancora non è stato scoperto, ma come trattamenti si fanno i soliti trattamenti che si fanno nel caso in cui il sist. immunitario si autoaggrava il suo stesso organismo, si danno delle "botte" di farmaci che deprimono le risposte immunitarie → il più famoso è il CORTISONE (è un antinfiammatorio, nel senso che c'è un'infiammazione che richiamo linfociti, il cortisone li emette, cioè ne blocca la produzione, e quindi l'infiammazione diminuisce), ma è una terapia che cura il sintomo, ma non cura la causa. Ancora oggi, non si sa la causa e quindi è difficile curarlo.

Slide 10) ORGANIZZAZIONE DELLA CORTECCIA CEREBRALE

Quando parliamo di EEG (segnale elettroencefalografico), tutta questa massa di neuroni di supporto, non ci interessa: ~~non~~ che non ci siano, non che anche loro diano segnale elettrico, ma non ci interessano nel senso che quando si parla di segnale EEG, ci si riferisce solitamente e in modo corretto, all'attività elettrica dei neuroni grigi, quelli funzionali. Quindi sappiamo fin da subito che nel segnale EEG, non ci sono informazioni metaboliche o infiammatorie, ma ci sono informazioni funzionali di quello che il soggetto sta facendo in una determinata condizione o in una det. zona → ~~quindi~~ la nostra corteccia cerebrale (dove sono posizionati maggiormente i neuroni grigi) e quindi sono all'esterno, ~~quindi~~ è spesso pochi mm ⇒ materia grigia. Mentre quello che c'è sotto, sono gli assoni che passano, e gli altri neuroni ^{non} funzionali e normalmente viene chiamata materia bianca.

alla nostra reazione. Allora tutte le volte che nel vedere una scena udire un suono, percepire un movimento, ecc., abbiamo una reazione, quella rosa nella zona frontale e si chiama anche corteccia associative (o zone associative). È chiaro che le diverse cose che vengono messe insieme, sono associate sulle base del nostro codice genetico, ma anche sulla base della nostra esperienza ⇒ parte è conoscenza innata, parte è esperienza.

SSS

La lobotomia è un intervento che si faceva nei malati psichiatrici e consisteva nell'asportazione del lobo frontale → togliendo le parte di associazione, cambiare il carattere.

I centri dell'udito sono dietro le orecchie e poi si ha il solco centrale (che abbiamo tutti) che divide la corteccia motoria dalle corteccie somatosensitiva: la corteccia motoria serve a tutti i movimenti volontari e a tutti quei movimenti che sono parte volontari e parte involontari (in es. il la respirazione o le motilità oculari → mettere a fuoco automaticamente, spostare gli occhi, ... ⇒ sono movimenti parte vol. e parte invol.).

La corteccia somatosensitiva invece fa la parte di acquisizione di tutti quegli stimoli che non sono stimoli motori (il caldo, il freddo, la vibrazione, ecc...) → nell'uomo non sono particolarmente spinti, però ci sono. La memoria è un po' distribuita (vedere mnestica), ma ci sono altre zone in giro per la corteccia cerebrale dove vengono immagazzinati i ricordi (c'è la memoria vera e propria): reti particolari di neuroni che immagazzinano i ricordi.

Il cervello è plastico, i neuroni ^{cambiano,} ~~immovono~~ e ricrescono (e da stabilire...), però di sicuro ramificano in modo ≠ e ramificano anche velocemente: tutti gli stimoli lasciano un'impronta a livello del nostro cervello.

Unici nel regno animale, ci sono 2 cose del nostro cervello che lo differenziano da quello delle altre specie:

1) LA LATERALITÀ → noi abbiamo quasi sempre un emisfero dominante

e richiede impegno, tanto più il numero di neuroni è grande. Quindi abbiamo un'area dedicata alle funzioni visive, che è enorme in area occipitale (prende quasi tutta la zona occipitale, perché dobbiamo vedere info tattiche, capire i colori, capire le forme, vedere le info in movimento e predire movimenti, pericoli ecc. ⇒ ci vuole un impegno enorme e poi in tempo-reale).

⇒ L'organo è eseguito più grande, quando il numero di neuroni coinvolti per comandare quell'organo, è maggiore. Per es., spendiamo moltissimi neuroni solo per comandare il pollice → è importantissimo per noi perché ci consente la presa di precisione. Lo stesso, con tutti i suoi gradi di libertà, ovviamente impiega molti neuroni; molto meno tutta la parte del corpo che ha dei movimenti limitati (il braccio, ginocchio, anca, caviglia ecc...) e inoltre abbiamo un'enorme muscolatura facciale → per comandare i muscoli del volto (compresi i muscoli oculari, delle palpebre e tutta la parte di sorghito, lingua, masticazione, deglutizione), impieghiamo una quantità di neuroni molto grande. Implicitamente, il fatto di dover dedicare un'area molto più grande per comandare un determinato organo, è un pro e un contro perché se l'area è più grande, nel caso di danno cerebrale, è più facile che quell'area venga danneggiata ⇒ succede che nel caso di danno cerebrale (es. ictus) in cui una parte di neuroni muore, capita per primo caso che, a seconda che prende l'emisfero dx o sx, si hanno problemi all'emicorpo controlaterale; poi è molto difficile che non riesca più a comandare ad es. il pollice o la caviglia. Invece è molto più facile avere problemi all'arto superiore (area più estesa), oppure ai muscoli legati all'espressività o alla deglutizione.

La cosa peggio che può accadere è quella di perdere le motilità delle mani e perdere la mobilità delle facce e del collo.

Si chiama **MONOCULUS SENSITIVO** → è l'organo splinato sulle sue cortecce cerebrali. I responsabili della corteccia cerebrale, sono i neuroni funzionali (quelli gialli delle slide 5).

Si vede che quando ti parlo di prelievo EEG, non metto 4, 2 o 3 elettrodi, ma metto una matrice di elettrodi in modo da coprire l'attività elettrica di tutto l'encefalo.

Questi neuroni per la forma che hanno, sono chiamati neuroni piramidali, perché hanno un corpo cellulare vagamente piramidale.

Se oltre alla componente elettrica, volessi sentire la componente magnetica, devo mettere in opera quello che si chiama: magneto-encefalografia \rightarrow è una specie di cavo di eroto liquido che si mette in testa, con i sensori che devono essere a pochi kelvin, c'è un problema di rapporto $\frac{S}{N}$ povero, è complicato, quindi la magnetoencefalografia viene usata solo per scopi di ricerca nei laboratori ma non per scopi clinici. Mentre l'EEG ha una valenza clinica di tutti i giorni.

Slide 8) POTENZIALE D'AZIONE E SINAPSI

Una sinapsi è fatto così è rappresentato alle slide 8 \rightarrow c'è l'assone che arriva (quella gialla è la guaina mielinica), poi ad un certo punto c'è una parte scoperta, poi l'assone ha un allargamento, detto bottone sinaptico, e va ad attaccarsi al corpo cellulare (o anche all'assone) di un altro neurone e da origine a quelle che è la vera e propria sinapsi. In realtà non si toccano proprio, ma c'è uno spazio intersinaptico (tra l'assone che arriva e il bersaglio), dove vengono rilasciati i neurotrasmettitori, perché la mediazione che avviene lì, è una mediazione chimica. Quindi, sull'assone corrono potenziali elettrici, e questo potenziale elettrico causa un certo n° di cambiamenti nelle membrane dell'assone, per cui si aprono canali (calcio che entra) e danno origine ad una maggiore permeabilità della membrana, e alcune vescicole che contengono il neurotrasmettitore, all'interno della sinapsi, si fondono con la membrana del bottone sinaptico e si vede che quando si fondono e si fondono, rilasciano il neurotrasmettitore. Questo

è stato rilasciato il neurotrasmettitore, per non sprecare niente, queste vescicole intermediano di nuovo il neurotrasmettitore e si riformano fino a pieno \Rightarrow c'è un ciclo di recupero di questo neurotrasmettitore.

Però se ho la stessa sinapsi, che continua ad essere depolarizzata da un flusso ininterrotto di potenziali d'azione che arrivano, dopo un po' di neurotrasmettitore non ce n'è più e quindi quella sinapsi va in periodo refrattario e non produce più nulla. Questi fenomeni, a livello complessivo, sono collegati a quello che noi interpretiamo come fatica, abitudine, ecc... \Rightarrow sotto prova un sist. nervoso sempre allo stesso stimolo uguale, crea dei cambiamenti: un po' perché i neuroni rispondono in modo \neq perché si sono abituati; un po' perché alcune sinapsi non funzionano più come prima (e non funzionano più).
 \Rightarrow Il nostro cervello ha bisogno di variabilità, in modo da fare lavorare in tempi diversi, sinapsi di tipo \neq .

slide (3-4) UN PO' DI STORIA \rightarrow Vedi slide

soggetto che reagisce in modo diverso a stimoli esterni o che non riesce bene in alcune funzioni, ha semplicemente un livello di inibizione più elevato.

I neuroni piramidali sono solo eccitatori → hanno il compito di spingere una funzione. Tutti gli altri neuroni possono essere anche tutti inibitori (quelli che non sono piramidali).

Il corpo, di questi neuroni, ^{che si chiama soma,} ha un nucleo centrale, dal soma partono tutti i dendriti, ce n'è uno più lungo che è l'assone che alle fine va a fare tanti bottoni ~~ogni~~ sinaptici (quanti non si sa); per aumentare la velocità di conduzione, l'assone è coperto e protetto da una guaina mielinica interrotta → dove ci sono i buchi si chiamano nodi di Ranvier: per cui c'è una conduzione molto veloce perché si salta di buco in buco e quindi la velocità di conduzione, rispetto al normale tessuto nervoso, aumenta di 10-12 volte.

(Sua: 16) SISTEMA 10-20

Il segnale EEG si misura sempre con un prelievo di tipo singolo differenziale significa che si posizionano gli elettrodi in modo non invasivo (direttamente sullo scalpo), la presenza dei capelli non è un problema (basta mettere tra cute ed elettrodo una quantità di gel conduttivo sufficiente per garantire il passaggio degli ioni).

Lo schema più famoso per il posizionamento degli elettrodi è il SISTEMA 10-20 detto così perché il posizionamento degli elettrodi viene fatto considerando la distanza tra NASION e INION e quella laterale e la distanza tra gli elettrodi è il 10% o 20% di quella distanza.

C'è bisogno di un sistema di riferimento quando preleviamo un segnale EEG perché si è visto che sulle cortecce cerebrali, le funzioni sono distribuite in modo preciso: c'è una corteccia frontale, c'è una parte motoria, una parte somatosensitiva, uditive, della memoria, linguaggio, vista, ⇒ avendo misurato un segnale EEG e quando si vuole sapere quali

Quindi prendo i valori del periodogramma che stanno dentro la finestra e ne faccio la media e disegno un nuovo periodogramma dove, \forall valore di frequenza (cioè $\forall f_i$), c'è la media dei P-valori e quindi viene una cosa liscia (—). Facendo quindi la media, tutte le variabilità: cosuale se ne va e quindi ottengo un andamento liscio, che dipende da $P \rightarrow$ però non c'è una regola per scegliere il valore di P corretto e quindi il vero problema di questo periodogramma è che il valore di P impatta fortemente le prestazioni. Quando si è parlato di media, di tecnica dell'orecchio si è detto che la media è l'equivalente matematico di un filtro P-BASSO e ogni P-BASSO ha una $f_c \rightarrow$ in questo caso la f_c dipende da $P \Rightarrow$ se medio pochi valori, tolgo un po' di variabilità ma non tanta; se invece medio tanti valori, sbatto tutto la variabilità e quindi è come se filtrassi con un filtro P-BASSO o banda passante molto stretta. Però il valore numerico di P corretto non sappiamo quale è \Rightarrow non è un periodogramma facile da eledere, ma certamente si può dire matematicamente che il periodogramma di Daniell equivale al periodogramma semplice, convoluto con la funz. di transf. di un filtro P-BASSO:

$$\tilde{P}_D(f) = \tilde{P}_{X_s}(f) * H(f)$$

È dato che il filtro P-BASSO lo decido io con P , quello toglie un po' di variabilità.

Il periodogramma di Daniell è, inoltre, polarizzato allo stesso modo del periodogr. semplice perché dato che la polarizz. dipende esclusivamente dai lobi secondari della finestra implicita \rightarrow dato che la finestra non è stata toccata, allora la polarizz. non è cambiata.

Quindi il periodogramma di Daniell ha la stessa polarizz. del periodogr. semplice, ma è più consistente perché ha una variabilità ridotta grazie all'effetto P-BASSO che gli abbiamo applicato.

Questo di Daniell, è l'unico periodogr. che viene calcolato a partire

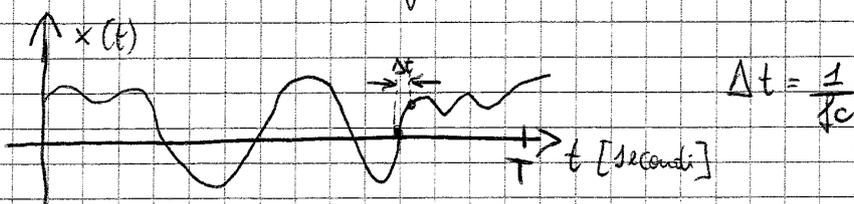
è quindi ~~facile~~ media, quello che era costante rimane costante e la variabilità casuale diminuisce.

Questo metodo funziona decisamente meglio perché è una specie di tecnica dell'overlapping rivisitata in chiave stima spettrale, dove quello che nell'overlapping era il segnale, in questo caso è lo spettro reale del segnale e quello che nell'overlapping era rumore che sovrapponeva il segnale, qui è la variabilità dello stimatore, che comunque è una variabile casuale.

Anche in questo metodo c'è ~~la~~ operazione di media e quindi c'è un equivalente P-BASSO che imponiamo alla variabilità casuale; ma chi definisce la f_c di questo P-BASSO? Il numero di bracci che abbiamo \rightarrow ma in realtà il n° di bracci che abbiamo è legato alla lunghezza del segnale e a quanto lungo facciamo ogni braccio perché noi dobbiamo ^{dire} che se abbiamo N campioni in totale e lo dividiamo in P segmenti, ciascuno di lunghezza D , ovviamente si deve avere che: $DP \leq N$.

Allora se voglio aumentare l'effetto P-BASSO, cioè ridurre sempre di più la variabilità, devo avere un valore di P sempre maggiore; però se il segnale è di lunghezza limitato, P più alto vuol dire segmenti più corti cioè vuol dire calcolare il periodogramma semplice di tanti pezzettini molto piccoli. Ma questo si può fare? Ovvero, posso ridurre arbitrariamente la lunghezza del pezzo di segnale su cui calcolo il periodogramma? Matematicamente lo posso fare, però bisogna ricordarsi che la trasformata di Fourier ci impone dei vincoli ~~arbitrari~~ arbitrario scemodi (in realtà inizia ad imporre il f_H del campionamento...) ma in generale l'accoppiata Nyquist più Fourier è micidiale.

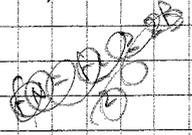
Supponiamo di avere un segnale che abbiamo ricostruito, campionato ad una f_c :



5.2.20) CARATTERISTICHE DEL PERIODOGRAMMA DI BARTLETT

La POLARIZZAZIONE del periodogramma di Bartlett è identica a quella del periodogramma semplice, dato che il periodogramma di Bartlett è basato su una media di periodogrammi semplici ed essendo il valore medio un operatore lineare, quando calcolo il valore atteso del periodogramma di Bartlett, è come calcolare la media del valore atteso del periodogramma semplice:

$$E\{\tilde{P}_B(f)\} = \frac{1}{P} \sum_p E\{\tilde{P}_{x,x}(f)\} = E\{\tilde{P}_{x,x}(f)\}$$



Inoltre non abbiamo fatto finestre \Rightarrow la polarizzazione cambia solo e sempre se metto una finestra.

La VARIANZA invece si riduce di un fattore P rispetto al periodogramma semplice (zi se già delle tecniche dell'overlapping):

$$\text{Var}\{\tilde{P}_B(f)\} \propto \frac{P_{x,x}^2(f)}{P}$$

5.2.27) METODO DI WELCH

Il periodogramma più usato di tutti, per la sua versatilità, è il periodogramma di Welch.

Welch fa una cosa molto semplice: copre che il fatto di avere segmenti di lunghezza uguale, uno dietro l'altro (non sovrapposti), molto spesso è limitante perché ho un'uguaglianza e lo taglio; ma per avere una risoluzione spettrale sufficiente, non posso tagliarlo troppo corto \Rightarrow il n° P di segmenti che ne derivano sono 3, 4 o 5, non sono mai tanti segmenti e quindi alla fine faccio un P-BASSO che taglia un po' di versatilità, ma non fa chissà cosa.

Allora, in molti casi pratici, la lunghezza del segnale D non ci consente di avere un valore di P molto elevato \Rightarrow Welch dice una cosa di questo tipo:

tutti i periodogrammi ottenuti, cioè:

f_c

$$\tilde{P}_w(f) = \frac{1}{P} \sum_{\Phi} \tilde{P}_{x_{i+\Phi}}^{(P)}(f) \quad (\text{vedi slide 58})$$

⇒ Welch propone la struttura dei segmenti, per ridurre le perturbazioni e la sovrapposizione dei segmenti, per aumentare il n° P (avendo N finito). Ma in realtà Welch, è più versatile perché posso anche scegliere di non aumentare P rispetto a Bartlett, ma di aumentare D ⇒ automaticamente ho lo stesso n° P di segmenti, ma ho segmenti più lunghi rispetto a Bartlett e quindi ho una risoluzione spettrale migliore.

Questo metodo di Welch ti pare di più, ma ci sono più cose da scegliere rispetto al periodogramma di prima (n° segmenti, lunghezza segmento, overlap e finestra).

In Matlab è una routine che ci permette di calcolare automaticamente Welch, cioè: dato un segn e specificata lunghezza di camp., lunghezza del segmento, finestra di stima che voglio usare e sovrapposizione, fa tutto lui → prende il segn, lo spezza nei segmenti opportuni (opportunosamente sovrapposti), finestra ogni segmento, calcola lo spettro di ogni segmento e fa la media di tutto (e restituisce questa media).

Quindi è molto comodo...

La funzione di Matlab si chiama: P Welch → in realtà P Welch non serve per calcolare solo il periodogramma di Welch, ma variando opportunamente i parametri d'ingresso, si può calcolare qualunque periodogramma, tranne Daniell.

Se per es. voglio calcolare un periodogramma dove i segmenti sono sovrapposti tra di loro, ma non finestra → questo periodogramma non ho nome, ma lo posso fare usando Bartlett finestra, cioè dividendo tutti segmenti consecutivi (non sovrapposti), finestra ogni segmento e calcolare lo

di vettori ortogonali che sono gli esponenziali di Fourier \rightarrow normalmente si ha un sistema lineare che ha esattamente tante condizioni, quante incognite (pochi $\approx N$, dove N è il numero di elementi del segnale).

\Rightarrow Quella che noi conosciamo come trasform. di Fourier a tempo discreto, si chiama DTFT. Quella che noi normalmente usiamo è la FFT (Fast Fourier Transform) \rightarrow la differenza è che c'è un algoritmo di calcolo particolarmente efficiente e vantaggioso per calcolare i coeff. che però si può utilizzare solo quando N è una potenza di 2! La differenza una volta era molto più accentratrice (DTFT più lenta, FFT più veloce), ma oggi (grazie alle potenze dei calcolatori) questa differenza non è così elevata \rightarrow anche se uso la DTFT invece che la FFT, su un n° di punti che non è quindi una potenza di 2, non cambiano significativamente le cose. Ma c'è una condizione: ciò vale quando devo fare 1 trasform. di Fourier; si vedrà però ~~che~~ alla fine del corso che 7 tecniche particolari (che si usano quando il segnale non è stazionario) che richiedono di fare un n° di trasform. di Fourier proporzionale al n° di campioni del segnale \Rightarrow se ho un segnale fatto da 1000 campioni, sono 1000 trasform. di Fourier; e se la DTFT ci mette il 25%-30% di tempo in più rispetto alle FFT, in alcuni casi non si riesce a terminare l'esecuzione in 3 ore! \Rightarrow bisogna abituarsi ad avere n° di punti ~~che~~ in frequenza ~~che~~ siano sempre potenze di 2!

Per le f. di campionamento che abbiamo e n° di punti, scegliamo fra 256 e 1024 o 2048, non ci sono tante potenze di 2 che consideriamo \rightarrow Quindi NFFT serve per stabilire quanti punti voglio in frequenza.

• $f_c \rightarrow$ freq. di campionamento del segnale

Infine gli ultimi parametri W e O sono i veri parametri che determinano quale procedimento voglio calcolare (sempre però tutti quanti, tecnica Donnell!): finestra e overlap!

Es. 13-14) CARATTERISTICHE DEL SEGNALE EEG

Il segnale EEG è un segnale casuale, perché quando misuro una qualunque degli elettrodi, e si fa la differenza tra il potenziale totale che sente questo elettrodo e un potenziale di riferimento, sto sentendo l'attività di un numero enorme di neuroni \Rightarrow ogni elettrodo sente una quantità di potenziali d'azione enormi contemporaneamente e tutti si sommano l'uno con l'altro e quindi statisticamente, questo è un segnale casuale, perché non si può sapere quanti sono i potenziali d'azione, come si sommano, ecc. Quindi i segnali che sono generati da fonti sorgenti, oppure delle quali teoricamente fa quello che vuole rispetto all'altro, danno origine a quello che si chiama segnali di interferenza. Questi segnali di interferenza, dal nostro punto di vista, sono sempre assimilabili a processi casuali \Rightarrow l'EEG è certamente sempre un processo casuale.

• AMPIEZZA \rightarrow l'ampiezza è variabile: si va da un minimo di una decina e un massimo di $\sim 500 \mu V$ (non è un segnale molto intenso, ci sono molti altri segnali del corpo umano più intensi, ad es. l'ECG si misura anche mettendoci un elettrodo alla caviglia --). I clinici distinguono l'ampiezza del segnale EEG in basso, media e alta, cioè se l'ampiezza è $<$ di $30 \mu V$ l'EEG ha un'ampiezza bassa; se è compresa tra $30 - 70 \mu V$ l'ampiezza è media; se è $>$ di $70 \mu V$ è considerata alta (considerando però un segnale EEG NORMALE \rightarrow vuol dire: non proveniente da stati patologici, in particolare non proveniente da un soggetto epilettico, che è l'alterazione più grande che c'è nella struttura del segnale EEG).

• MORFOLOGIA \rightarrow la morfologia ci dice come un segnale, ad una det. frequenza si manifesta sul nostro sistema di acquisizione.

Le morfologie possono essere POLICORFA o MONOCORFA \rightarrow vuol dire che se ho un segnale EEG che è formato da una successione di onde o di potenziali

• SINCRONIA/ASINCRONIA → se l'asimmetria/simmetria ci dice il dove, la sincronia/asincronia ci dice il quando: infatti ci possono essere eventi sul μ EEG che si presentano con le stesse tempistiche contemporaneamente sui 2 lotti \Rightarrow si dice che sono μ SINCRONI (o danno origine ad μ sincroni), altrimenti sono ASINCRONI.

Quindi: le caratteristiche (a parte l'ampiezza e la topografia), dicono come è fatto (la morfologia), dove (la simmetria) e quando (la sincronia) e ~~da~~ un medico guarda queste cose quando ha un insieme di μ che formano il suo EEG.

Slide 15) EEG COME PROCESSO CASUALE

Quando abbiamo un μ EEG, otteniamo un certo n° di tracciati. Nel grafico sono rappresentati 4 canali, in orizzontale c'è il tempo in secondi e ci sono rappresentate 2 cose particolari:

• la 1° cosa particolare è l'artefatto tipico che c'è sul μ EEG: l'EEG soffre pochissimo dell'interferenza di rete, soffre pochissimo della contaminazione da altri μ biologici (quasi EMG o ECG), però ha un artefatto tipico che è l'attività dei muscoli oculari e in particolare lo battere delle palpebre. Soprattutto sui canali frontali (quelli più vicini agli occhi) quando apro e chiudo gli occhi sto attivando dei muscoli che sono innestati sotto lo scalpo e quindi quell'elettrodo sente direttamente l'attività elettrica di quel muscolo \Rightarrow in questo caso il soggetto ha aperto e chiuso gli occhi e si vede come sul canale frontale questo artefatto sia molto evidente e man mano è sempre ^{meno} evidente, allontanandosi dalle zone frontali.

Quindi l'artefatto classico del μ EEG è il battito delle ciglia, ma non è presente su tutti i canali, ma solo in quelli più in zona frontale. Come si rimuove? Ci sono molte tecniche: alcune sono banali e fanno una specie di filtraggio \rightarrow tramite filtraggio

dominante (o una banda di freq. dominante).

Guardando bene questo spg EEG, si è un spg coseno, ma in alcuni tratti sembra proprio una sinusoidale, cioè ha un andamento regolare.

Le lettere alfa e beta indicano valori di bande di frequenza diverse e queste bande sono correlate allo stato del soggetto. Allora se si prende come il fatto lo spettro di un spg EEG, si trovano tutte le bande di frequenza: delta, teta, alfa, beta, gamma.

(Slide 17) CARATTERISTICHE SPETTRALI

La tabella indica: quali sono i tipi di bande (o ritmi) che si possono trovare nel spg EEG; il range di frequenze; l'ampiezza (in μV); lo stato fisiologico del soggetto, associato prevalentemente a queste bande. Vol dire che, se il soggetto si trova in un certo stato, l'EEG prevalente ha uno spettro che è concentrato in quelle bande; è sempre un spg coseno, ma ha uno spettro che cambia e si concentra in frequenze che sono un po' le misure di quello che è lo stato funzionale del soggetto stesso. Queste bande sono (in ordine alfabetico in ordine temporale in cui sono state scoperte).

• BANDA DELTA \rightarrow banda a freq. più bassa (0.5 - 3 Hz) ed è associata a condizioni patologiche, o a sonno molto profondo (coma)

\rightarrow (le cellule neurali si riorganizzano)

È un ritmo molto lento (oscillazioni lente e molto ampie) \rightarrow si nota che al crescere della frequenza, decresce l'ampiezza del spg EEG

• BANDA TETA \rightarrow (da 3 - 7 Hz) ha un'ampiezza leggermente più bassa ed è sempre associata ad uno stato di sonno (non profondissimo)

BANDA ALFA \rightarrow (da 8 - 13 Hz) è la banda del rilassamento mentale (è la fase della veglia al sonno).

~~Le onde alfa sono le onde di Berger, le prime trovate~~
Le onde alfa sono le onde di Berger, le prime trovate

per sapere cosa succede e gli elettrodi relativi (collegati alle slide) sono quelli della zona centrale:

• (b) → sono leggeri. C'è una X: durante questa fase di sonno leggero, al paziente è stato fornito uno stimolo e si è andato a vedere se il soggetto rispondeva o meno ⇒ in corrispondenza di questo stimolo, si vede un picco su qualche canale. Questo vuol dire che il soggetto era ancora in allerta, rispondeva ancora allo stimolo esterno, anche se effettivamente stava dormendo.

• (c) → sempre sono leggeri (ritmi beta dominanti)

• (d) → sono più profondo

~~• (e) ed (f)~~ • (e) ed (f) sono le ultime fasi del sonno → si vede che sono più profonde, perché le onde diventano sempre più lente e più ampie.

→ Se prendo un EEG, posso capire un soggetto in che fase del sonno è, proprio uno degli usi che si fanno del pg EEG, è quello di studiare le patologie del sonno. Sulle patologie del sonno sono stati scritti libri di tutti i tipi.

(Bianco) LE FASI DEL SONNO

Il sonno è importantissimo, perché è il momento in cui il sistema nervoso centrale, riorganizza se stesso, fa un backup. Se non dà il tempo al sistema nervoso centrale di riorganizzare l'informazione, non si fissano i ricordi, siamo meno pronti a rispondere agli stimoli e quindi il sonno più è profondo e più è efficace ed è importantissimo per il sistema nervoso centrale per organizzarsi se stesso.

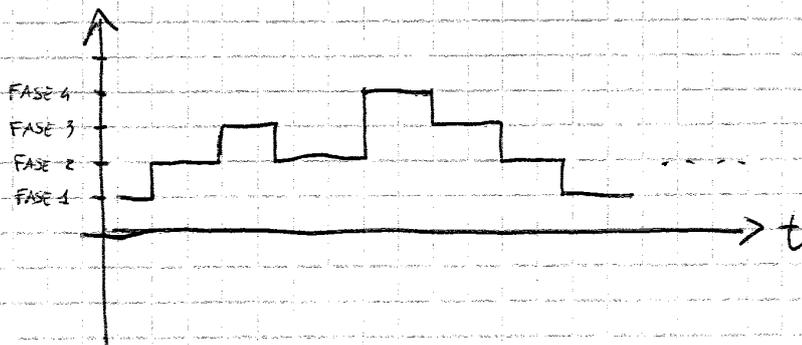
Le cliniche del sonno sono centri specializzati per studiare le patologie del sonno: si prende il soggetto, lo si fa addormentare con un cerchietto per l'EEG (e anche un ECG, perché non si può sapere le patologie del paziente), un

Il bambino ha solo 2 fasi: veglio e in "coma". L'adulto non è in grado di passare dalla veglia al sonno profondo, ma attraverso le altre fasi del sonno. Con l'aumentare dell'età, il sonno diventa sempre più superficiale → si dorme sempre meno e si va sempre meno in fase ~~profonda~~ di sonno profondo, che però è quello più utile.

Nel dettaglio:

- fase 1 → sonno leggero, c'è un rallentamento dell'attività muscolare e occasionalmente ci sono degli spasmi muscolari.
- fase 2 → c'è un rallentamento del ritmo respiratorio e del ritmo cardiaco. Comincia a diminuire la temperatura corporea.
- fase 3 → è la prima fase dove inizia il sonno profondo e qui cominciano le onde S. Quindi fino alla fase 2, siamo tre onde theta; le onde S iniziano in fase 3.
- fase 4 → sonno molto profondo, il respiro diventa ritmico (quindi è totalmente involontario), l'attività muscolare è molto limitata e l'EEG è praticamente fatto tutto da onde S.

In questa classificazione, ci si dà bene che le vere fasi del sonno, alle fine, sono la 2 e la 4, che sono il sonno (fase 2) e il sonno molto profondo (fase 4). La fase 1 e la fase 3 sono fasi di transizione, e ci sta poco tempo: la fase 1 è la transizione tra il sonno e la veglia, mentre la fase 3 è la transizione tra il sonno e il sonno molto profondo. Si può costruire un diagramma del sonno durante la notte, in cui si vede in che fase era il soggetto: ~~diagramma~~



colore la potenza che c'è in ciascuno di queste bande. Fare l'analisi spettrale e calcolare la potenza in ogni banda, sono oggetti del laboratorio 4.

OSS.

Il segnale EEG è un processo casuale stazionario in senso lato?

Se si considerano 6h di segnale EEG, certamente non è un segnale staz. in senso lato; se cambia il suo contributo spettrale in diverse frequenze, vuol dire che cambia la sua funzione di autocorrelaz. con il tempo \Rightarrow non è staz. in senso lato.

E allora come si fa l'analisi spettrale di questo segnale? Non si fa l'analisi spettrale di un brano di 6h, ma ne prendiamo dei pezzettini, tanto il cambiamento tra fase 1, fase 2, fase 3 ~~non~~ avviene istantaneamente, ma c'è una transizione tra una fase e l'altra \Rightarrow si prendono dei pezzettini in cui si è abbastanza sicuri che in quella finestra il segnale EEG sia assimilabile ad un processo casuale WSS (cioè staz. almeno in senso lato). Questi pezzettini quanto sono lunghi?

~~Se ho 2 minuti del segnale EEG, che risoluzione spettrale teorica si ha (trascurando il fattore di qualità)?~~

$$\Delta f = \frac{1}{120} = 8,3 \text{ mHz} \rightarrow \text{dato che c'è un interesse distinguere tra}$$

bande diverse, una $\Delta f = 8,3 \text{ mHz}$, serve? È troppa, perché nelle bande tutte, ad es., che va da 30 a 7 Hz (che è fatta da 4 Hz), ogni Hz \Rightarrow è rappresentato da ~ 100 punti \Rightarrow vuol dire che le bande fatte e rappresentate da 400 punti e il fatto di prendere 2 minuti (120 secondi), è un metodo eccellente per introdurre una variabile enorme.

\Rightarrow Visto che tutte le bande più o meno sono distanziate di 4-5 Hz (tranne la beta), un valore di Δf che ci può andare bene è di 1 Hz (al minimo), ma anche 0,5 Hz. Ma se voglio $\Delta f = 0,5 \text{ Hz}$, quanto EEG mi serve prendere? 2 secondi! Se 2 secondi è rappresentabile supporre che il segnale EEG sia stazionario in senso lato? Sì!

Alcuni sostengono che c'è una rete di neuroni che commuove e lavora molto ed elettricamente causa la depolarizzazione di tutti i neuroni che gli stanno attorno e quindi da un focolaio si propaga l'attacco epilettico. Altri dicono che gli attacchi epilettici nascono in più punti diversi e si sincronizzano per qualche motivo e poi causano la disfunzionalità di tutta la corteccia cerebrale.

Quello che succede veramente, non si sa ancora, ma si sa come si tratta: psicofarmaci, che hanno il compito di alzare il livello delle reti inibitorie: infatti il soggetto epilettico ha un livello di inibizione più basso degli altri e quindi basta poco per far scattare molta attività cerebrale, che però non fa bene.

La grosso sfida dell'epilessia è prevenirla, cioè tentare di capire quando il soggetto sta osservando l'attacco epilettico → questo si può fare, infatti esistono delle tecniche di analisi del segnale non lineari che sono in grado (più o meno) di prevedere l'occorrenza di un attacco epilettico. Nella pratica hanno un'utilità quasi nulla, ma nei soggetti che sono ricoverati (che hanno tanti attacchi o che sono farmaco-resistenti), lo studio del segnale EEG per la prevenzione dell'attacco epilettico, può aiutare a sapere se una terapia farmacologica funziona meglio di un'altra, se mediamente il suo livello di inibizione è aumentato, ecc...

Slide 22) ONDE MU

Sono molto usate oggi da chi si occupa di interfaccia uomo-macchine. Le onde mu sono delle onde alla frequenza di 7.5-12 Hz e quindi si sovrappongono alla banda alfa; solo che queste onde sono presenti solo nella corteccia motoria: sono onde specifiche della corteccia motoria. Se noi registriamo il segnale EEG dalla corteccia motoria, queste onde ⁱⁿ le troviamo, e meno che il soggetto non stia usando la corteccia motoria per volere un movimento. Quindi l'attivazione della corteccia motoria per

e c'è una certa attività cerebrale, nella stessa area, ~~se~~ se lo scimmio vede fare lo stesso azione di prendere la banana, c'è la stessa attività cerebrale \Rightarrow stesso punto e stessa attività (anche se lo scimmio è fermo). Sono neuroni che si attivano ^{tra} quando il soggetto fa un'azione motoria, sia quando vede compiere un'azione motoria e il livello di similitudine varia a seconda del tipo di attività: per attività molto semplici è facile, per attività motorie complesse, il livello di similitudine diventa via via minore.

Sono presenti anche nell'uomo questi neuroni, anche se nell'uomo hanno un funzionamento che si suppone differente: si pensa ci sono neuroni specchio motori, ma anche neuroni specchio somato-sensitivi, che sono quelli che ci danno l'empatia (se vedo uno che piange, divento triste).

Succede che i neuroni specchio, sulle scimmie, si attivano non tanto in corrispondenza di un'azione, ma quando il sist. nervoso centrale intuisce quale è l'azione da fare \Rightarrow anche qui è importante l'intenzione che precede l'azione di fare se \Rightarrow se si fa un esperimento (slide 24): lo scimmio prende una fetta di melone e lo mangia e c'è una certa attività, con picco nel momento in cui afferra la fetta di melone (1), nell'esperimento (2) lo scimmio prende una fetta di melone e lo mette in un contenitore; negli esperimenti (3) e (4), le azioni sono le stesse, solo che lo scimmio prende un umano che lo fa e succede che anche in questi 2 esperimenti, c'è il picco dell'attività quando l'umano afferra il melone e lo scimmio capisce che lo mangia. ~~È diverso~~ è diverso quando l'umano afferra il melone e lo pone in un contenitore o lo mangia: in esse caso c'è un'attività molto simile a quando lo fa lo scimmio direttamente, nell'altro caso (quando l'umano lo prende), non può sapere se lo mangia o meno e quindi l'attività viene dopo, cioè il picco di attività viene dopo l'inizio dell'attività dell'umano.

05/06/2015

ULTIMA LEZIONE

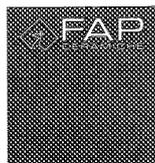
riepilogo..

La scorsa volta abbiamo visto come abbiamo chiamato complessità fatto il segnale nel dominio del tempo e la frequenza → ha più

Abbiamo visto con un esempio

modo per misurare, quella che chiamata di Fourier o da come è una cosa che sta a metà tra serie numerica si ripete.

spettro così,



quando un segnale può avere

l'entropia è elevata, perché abbiamo poche ricorrenze sul numero di frequenze molto elevate. Vuol dire che i segnali più complessi, hanno spettri bassi e larghi; mentre i segnali meno complessi, hanno spettri con picchi più alti e banda più stretta.

Avevamo detto che si poteva definire anche come la definisce Renyi, tanto cambia poco:

$$RE = -\log \left(\sum_f p_{xx}^2[f] \right)$$

il discorso alla fine è che questa misura di entropia, è sempre legata alla trasformata di Fourier e quindi per noi soffre di tutti i problemi di stima spettrale (perché p_{xx} dovrebbe teoricamente essere lo spettro del segnale che non abbiamo e sostituiamo con la stima dello spettro del segnale: periodogramma, correlogramma, ecc..).

Slide 6÷11 - Approximate Entropy - Example - Sample Entropy

Vediamo una definizione di entropia, che sia chiamata *entropia approssimata (ApEn)*, che prescinde dalla trasformata di Fourier. In modo più coerente rispetto alla definizione di complessità, va a vedere quali sono le variazioni all'interno di una serie numerica (cioè di un segnale campionato), che sono più o meno predicibili → cioè va a vedere all'interno di una serie di segnali, quanto i campioni di questo segnale, ricorrono (a diverse distanze di tempo).

Quello che noi chiamiamo predicibilità o ricorrenza, non ha niente a che vedere con la correlazione: non è una misura di correlazione. È una misura di similarità o somiglianza, che in questo caso che vedremo, è effettuata solamente sulla base di analisi della serie temporale: niente Fourier, niente correlazioni, ma solo analisi dei numeri nel dominio del tempo.

Questa è una delle tante definizioni di entropia → questa è una definizione che per segnali biologici, si è visto che funziona abbastanza bene.

Dice che intuitivamente si può pensare che, la presenza di pattern di fluttuazioni, che si ripetono nella serie temporale di un segnale, renda questo segnale più predicibile, rispetto ad una serie numerica in cui questi pattern sono assenti (pur essendo tutti e due segnali casuali).

⇒ *L'ApEn* esprime la verosimiglianza logaritmica che, all'interno di un segnale di lunghezza N , ci siano delle sequenze che si ripetono uguali per una distanza pari almeno ad r essendo sequenze di d punti e che si ripetono anche quando la sequenza diventa lunga $d+1$ punti.

→ costruisco tanti vettori (o sequenze), che però non contengono né Fourier né niente, ma sono gli stessi campioni del segnale nel dominio del tempo, non cambia nulla.

ESEMPIO → abbiamo un segnale periodico, fatto da questa sequenza numerica:

$$x = \{61, 62, 63, 64, 65, 61, 62, 63, 64, 65 \dots\}$$

Abbiamo $N=50$ (campioni), e ovviamente è un segnale periodico di 5 campioni e quindi vuol dire che in 50 campioni ci sono 10 periodi.

Imponiamo $d=5$ (esattamente pari al periodo di questo segnale) e $r=2 \Rightarrow$ se andiamo a costruire i vettori numerici, $X(1)$ sono i primi 5 numeri, $X(2)$ inizia dal secondo, e così via:

$$X(1) = \{61, 62, 63, 64, 65\}$$

$$X(2) = \{62, 63, 64, 65, 61\}$$

$$X(3) = \{63, 64, 65, 61, 62\}$$

$$X(4) = \{64, 65, 61, 62, 63\}$$

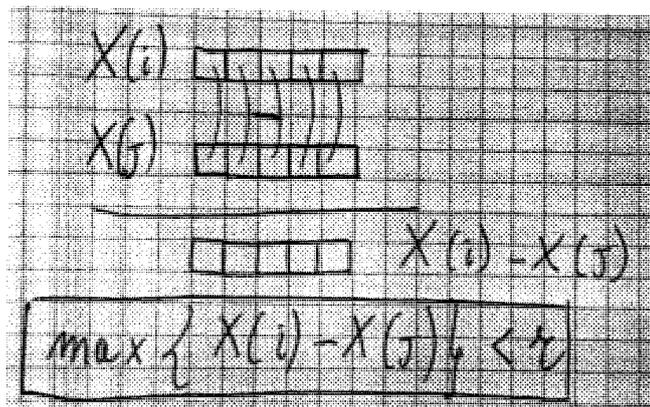
$$X(5) = \{65, 61, 62, 63, 64\}$$

...

ne costruisco $N-d+1 \Rightarrow$ con $N=50$ e $d=5$, il numero totale di vettori è 46.

Adesso devo andare a contare quante di queste 46 sequenze, posso considerare simili tra di loro. Quando, per me, due sequenze sono simili?

Supponiamo di avere la sequenze i -esima $X(i)$ e la j -esima $X(j) \rightarrow$ come faccio a sapere se queste due sequenze sono simili? Faccio la sottrazione, elemento per elemento, di quei due vettori e ottengo il vettore $X(i) - X(j)$. Allora dico che queste due sequenze sono simili, se il massimo di $X(i) - X(j)$ (cioè la massima delle differenze), è minore di r (il secondo parametro che mi sono imposto prima):



⇒ se facendo la differenza, elemento per elemento, scopro che la massima differenza (tra quelle cinque) è minore di r , allora dico che quelle due sequenze sono simili.

Vuol dire che: preso quel pezzo di segnale, a qualunque distanza, quel pattern di numeri si è ripetuto ragionevolmente uguale → a qualunque distanza, perché poi lo devo fare per tutte le sequenze. Quindi devo confrontare ogni sequenza i -esima con tutte le altre sequenze.

L'entropia approssimata abbiamo detto che è definita come la verosimiglianza logaritmica che, all'interno di un segnale di lunghezza N , ci siano delle sequenze che si ripetono uguali per una distanza pari almeno ad r essendo sequenze di d punti e che si ripetono anche quando la sequenza diventa lunga $d+1$ punti, cioè un campione in più. Quindi ci dice che l'entropia bisogna misurarla in modo corretto: non basta che si ripeta data una certa lunghezza, ma si deve ripetere anche ragionevolmente se allungo d di un campione, proprio per evitare periodicità di segnali di questo tipo.

⇒ prima ho calcolato C per un valore pari a 5 (con una tolleranza di 2), adesso devo calcolare C anche quando la lunghezza dei segmenti è pari a 6, cioè $d+1$:

$$C^6 = \frac{45 \cdot \frac{9}{45}}{45} = 0,2$$

Siccome 6, non è uguale a 1 modulo 5, allora il numero di volte in cui la sequenza è simile, è solo 9 (non è 10) e questa volta vale per tutte le sequenze => ci sono 45 sequenze che si ripetono, uguali, 9 volte su 45.

Adesso, l'entropia approssimata è il logaritmo del rapporto tra C^5 e C^6 :

$$ApEn = \ln\left(\frac{C^5}{C^6}\right) = 0.00189$$

⇒ questa è un'entropia molto bassa (infatti è un segnale periodico, ed è quello che ci aspettavamo); se invece lo facessimo per un processo casuale, troveremmo valori di ripetizione molto diversi, e questo logaritmo diventerebbe molto più alto, cioè salirebbe il valore dell'entropia.

Definita in questo modo, al minimo, l'entropia può valere 0 (nel caso in cui C^5 e C^6 sono uguali); C^5 e C^6 non possono mai valere 0, perché ogni sequenza è sempre ad una distanza minore di r da se stessa => al minimo C^5 e C^6 valgono 1. Esiste un valore superiore di questa entropia? C è definito, per ogni sequenza, come il rapporto fra il numero di sequenze che hanno la distanza corretta e il numero totale delle sequenze. Quindi se suppongo un segnale costante, ovviamente avrei che tutte le sequenze hanno la distanza corretta e quindi C varrebbe 1 (e quindi torniamo nel caso precedente). Al contrario, abbiamo detto che al minimo C può valere 1 → però questo 1 andrebbe diviso per il numero delle sequenze, che a priori non conosciamo perché dipende dalla lunghezza del segnale e da d => il valore di C^d massimo non possiamo saperlo a priori.

Quindi, se non so quali possono essere i valori massimi, posso solo dire che $ApEn$ andrà da 0 in su: più è alta, più complesso è il segnale; però non ho un limite di riferimento preciso (anche se nella pratica si scopre che poi avere valori di entropia superiori all'unità, è abbastanza complicato).

Da questa scrittura:

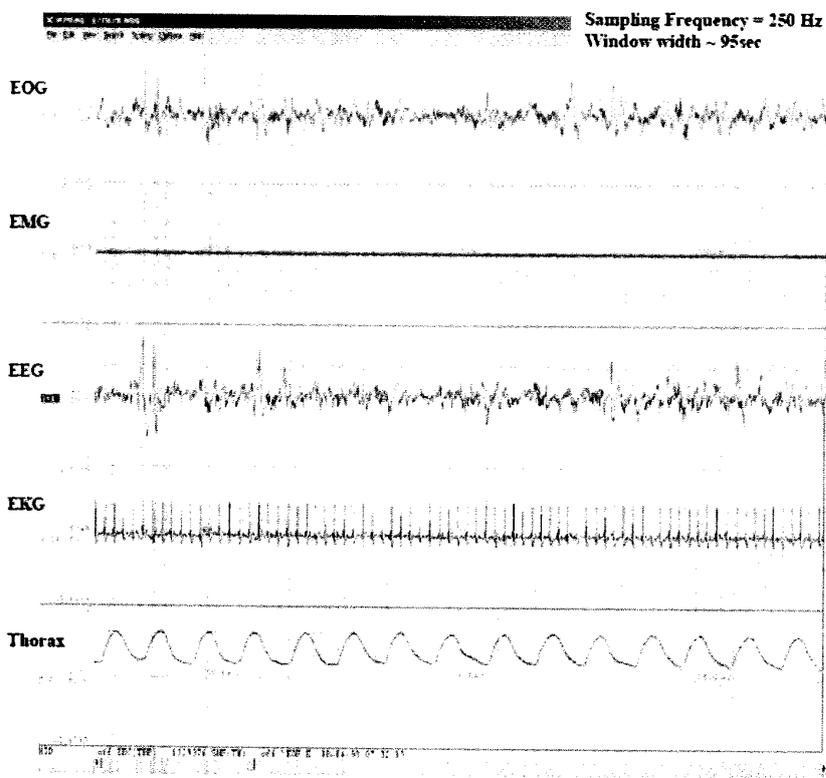
$$ApEn(d, r, N) = \phi^d(r) - \phi^{d+1}(r)$$

$$\phi^d(r) = \frac{1}{(N-d+1)} \sum_i \ln(C_i^d(r))$$

si può ridefinire l'entropia approssimata in questo modo qua (che è quella che abbiamo usato prima):

$$ApEn(d, r, N) = \frac{1}{N-d+1} \sum_i -\ln\left(\frac{C_i^{d+1}(r)}{C_i^d(r)}\right)$$

Quello è il classico tracciato di quella che si chiama *polisonnografia*:



Quando un soggetto va in clinica per fare l'analisi del sonno, gli misurano molti parametri; qui ci sono:

- l'elettrooculogramma (*EOG*);
- l'*EMG* (tipicamente di un muscolo articolare delle gambe, per il motivo che i soggetti che hanno problemi, la prima cosa che muovono sono le gambe);
- l'*EEG* (dove si vede bene che tutte le volte in cui l'*EOG* fa vedere un picco di attività, questo picco si ritrova anche sull'*EEG* → quello che abbiamo detto è l'artefatto da *eyes blinking*, cioè da sbattimento delle palpebre: soprattutto i canali più frontali sentono il segnale EMG dei muscoli oculari quando si sbattono le ciglia);
- l'*EKG* (o *EKG*);
- il tracciato della respirazione (*Thorax*) → è l'impedenziometria toracica, che dice quanto il torace si dilata e si comprime, da cui si estrae il ritmo respiratorio.

Si può osservare come l'entropia, all'aumentare della profondità del sonno, diminuisce (a significare che il tracciato EEG è sempre più semplice), mentre quando si va in fase REM, risale.

Abbiamo detto che il cervello anche quando dorme non è che non fa niente (non fa cose relative al funzionamento dei neuroni corticali, perché si sta occupando di fare altre cose) => siccome con l'EEG noi sentiamo solo l'attività dei neuroni corticali, man mano che il sonno è sempre più profondo, quell'attività è sempre minore, sempre più ripetibile, sempre più lenta. Poi nella fase REM si sogna e quindi si spende di più, nel senso che i neuroni si riattivano perché fanno sentire sensazioni => si ha uno stimolo maggiore e questo si vede dall'entropia che aumenta.

Quindi l'entropia descrive già da sola, la profondità del sonno.

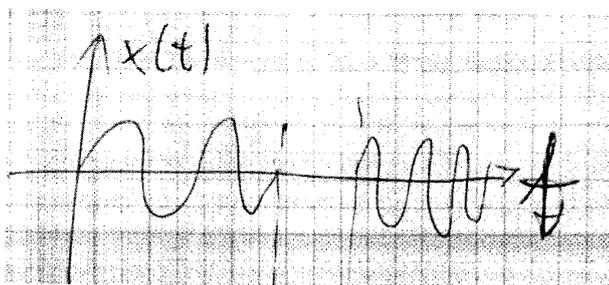
Se ci dovesse chiedere, data una serie numerica, di calcolare un certo numero di entropie (ne conosco 4: spettrale, di Renyi, approssimata e campionaria), so come si fa → sono tutte semplici e sono tutte buoni descrittori della complessità strutturale di un segnale che noi poi riferiamo alla complessità del sistema fisiologico che l'ha prodotto.

Il laboratorio dice:

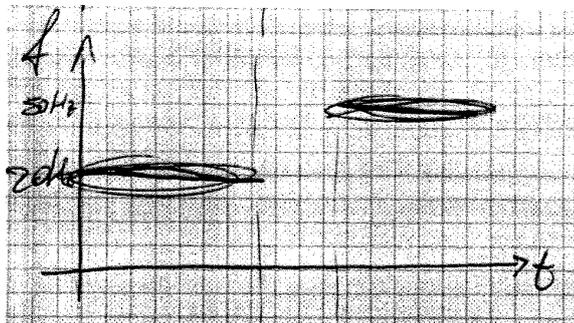
1. Ricavare la rappresentazione tempo-frequenza tramite trasformate di Wigner-Ville e di Choi-Williams dei segnali sintetici descritti di seguito, assumendo per tutti i segnali una frequenza di campionamento pari a 256 Hz e un supporto temporale pari a 1 s.
 - somma di due sinusoidi a 20 Hz e 50 Hz rispettivamente con supporto $0 \div 0.4$ s la prima e $0.6 \div 1$ s la seconda

Vuol dire che ho un segnale lungo 1 secondo (quindi un vettore di 256 campioni, dato che è 1 secondo a 256Hz) e tutti i campioni che vanno da 0 a 0.4 secondi, contengono una sinusoide a 20Hz; poi ho un pezzo in cui non ho niente, e poi tutti i campioni che vanno da 0.6 a 1 secondo, contengono una sinusoide a 50Hz.

→ vuol dire che il segnale è fatto così:

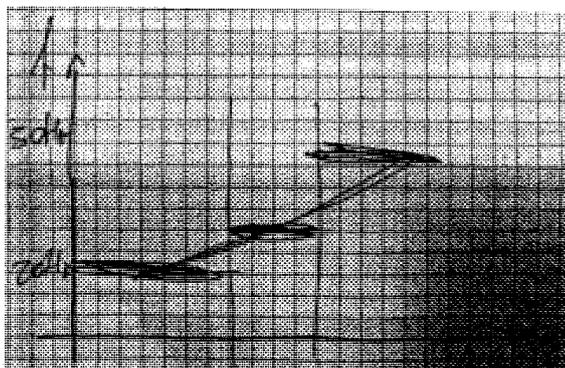


e quindi mi aspetto questa rappresentazione t-f:



cioè in mezzo niente, sulla sinistra una retta a 20Hz, e a dx mi aspetto il contributo a 50Hz.

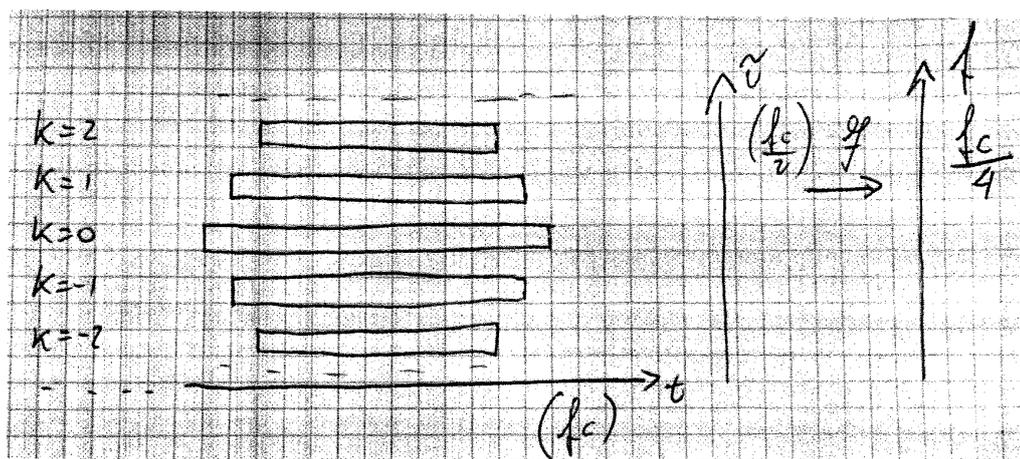
MA NON È COSÌ: se calcolo la distribuzione di WV, mi ritrovo i termini interferenti e abbiamo detto che questi termini si localizzano a metà, sia nel tempo che in frequenza e quindi sulla congiungente tra il termine a 20Hz e quello a 50Hz, ci saranno i termini interferenti:



- quando $k=2$, analogamente, ho un prodotto che è di 2 campioni più corto del caso precedente.
E così via...
- ⇒ se prendo tutti i prodotti di autocorrelazione e li metto in pila uno sull'altro, ottengo una piramide che si stringe di 2 campioni per ogni ritardo in più e quindi vuol dire che ci perdo un campione ogni volta.

Quindi si ha il numero totale di righe che ha questa matrice, non è quello che ci aspettiamo (imposto un certo ritardo), ma è la metà!! → imposto un certo valore di ritardo ($nt \mid ag$), non ho tutti i prodotti di autocorrelazione come mi aspettavo nel caso dell'autocorrelazione reale, ma ne ho la metà. Siccome però le righe di questa matrice corrispondono all'asse τ , questo asse τ è quello che poi darà origine all'asse della frequenza (tramite trasformata di Fourier).

Allora se sull'asse τ ho la metà dei campioni, è come se avessi una frequenza di campionamento che è la metà => quando vado in frequenza, la massima frequenza rappresentabile non è più $\frac{f_c}{2}$, ma è la frequenza di campionamento dell'asse $\frac{\tau}{2}$; ma la frequenza di campionamento dell'asse τ , è la metà della frequenza di campionamento. Quindi in totale, il limite superiore dell'asse τ diventerà: $\frac{f_c}{4}$ e non $\frac{f_c}{2}$.



Quindi si ha un numero di prodotti di correlazione metà rispetto a quello che dovrei avere, e quindi è come se avessi l'asse verticale del ritardo sotto-campionato di un fattore 2.

→ se f_c è la frequenza di campionamento dell'asse temporale, l'asse del ritardo ha una frequenza di campionamento che è la metà. Essendo che è l'asse del ritardo che dà origine all'asse della frequenza, ovviamente la massima frequenza è la metà della frequenza di campionamento dell'asse τ , cioè $\frac{f_c}{4} \Rightarrow \frac{f_c}{4}$ se $f_c=256 \text{ Hz}$, fa $64\text{Hz} \rightarrow 80\text{Hz}$ è superiore a 64Hz e quindi vuol dire che si può rappresentare tra 0 e 64Hz , ma si ha un contributo che è sopra $64\text{Hz} \Rightarrow$ fa aliasing e, come tutti gli aliasing, rientra dall'altra parte e quindi rientrando dall'altra parte, sposta la posizione degli interferenti e quindi succede un macello.

⇒ questo è il caso classico dato apposta per vedere cosa succede.

Posso dire: "quando uso queste trasformata t-f devo campionare ad una frequenza, che invece di essere doppia, deve essere 4 volte la massima frequenza contenuta nel segnale, per non avere aliasing?" Ovviamente no, perché esistono anche implementazioni che preservano tutti i prodotti i correlazione e che quindi non determinano sotto-campionamento dell'asse verticale.

13/04/2015

CONTRAZIONI MUSCOLARI VOLONTARIE

Il meccanismo dietro la generazione del segnale è diverso a seconda che una contrazione muscolare avvenga volontariamente o perché stimolata dall'esterno. Nel caso di contrazioni volontarie abbiamo visto che c'è un insieme di motoneuroni che innervano le unità motorie, il SNC pilota ogni unità motoria indipendentemente dall'altra e quindi con una serie di elettrodi misuriamo una serie di unità motorie ciascuna delle quali ha un'attività disorganizzata. Come succedeva per il SNC con il segnale EEG, anche il segnale EMG è casuale, è un classico segnale di interferenza in cui si sommano nel tempo i contributi di diverse sorgenti.

In questo caso abbiamo la sommatoria dei così detti treni di potenziale d'azione di un'unità motoria che danno origine ad un processo casuale sui quali poi si innesta un rumore casuale, la funzione di filtro del sistema di prelievo e dà origine al segnale EMG osservato.

Abbiamo visto anche che c'è differenza tra le funzioni di trasferimento nel caso di prelievo invasivo e nel caso di prelievo di superficie, perché nel primo caso non c'è una funzione di trasferimento dei tessuti, che è un filtro passa basso, con frequenza di taglio e attenuazione sempre più forte all'aumentare dello spessore, per cui su soggetti particolarmente voluminosi o muscoli profondi il prelievo di superficie diventa difficile.

GENESI DEL SEGNALE EMG

Il segnale EMG è un classico segnale di interferenza in cui i treni di potenziale d'azione di unità motoria si sommano nel tempo.

MUAPT = treni di potenziale d'azione di unità motoria.

$$s(t, F) = \sum_{i=1}^P u_i(t, F)$$

Per noi il modello indicato va bene, ci dice che il segnale EMG s , quello che noi abbiamo, è pari alla sommatoria dei treni dei potenziali d'azione di unità motoria dove tutto dipende dal tempo ma anche da un altro parametro che è la forza di contrazione. Questo vuol dire che in prima battuta la forza di contrazione è il modo nel quale il SNC decide di reclutare le unità motoria, cioè quali, quante, come farle attivare.

Il SNC stima la forza di contrazione necessaria a compiere un movimento e attiva le unità motorie in numero tipologia e intensità coerente con la forza che bisogna esercitare. È il motivo per cui riusciamo ad avere con la stessa articolazione dei movimenti a forza molto bassa (movimenti di precisione) ma anche a forza molto più intensa. Quanto reclutarle e quanto farle attivare.

Il segnale EMG complessivo è uguale alla somma temporale dei treni di potenziale di unità motoria che vengono reclutati data una certa forza di contrazione. Quante unità motorie vengono attivate? Supponiamo un numero pari a P : il muscolo è fatto da tante unità motorie, se io ne attivo contemporaneamente di più la forza totale esercitata dal muscolo aumenta, perché più unità motorie si contraggono contemporaneamente. Dovremmo scrivere quindi che P è funzione della forza di contrazione F .

Supponiamo di mettere due elettrodi per effettuare un prelievo di segnale EMG di superficie su bicipite omerale e chiedo al soggetto di eseguire una contrazione volontaria, cioè flettere il gomito. Otterrò un segnale stazionario oppure no? No, perché dipende dalla forza di contrazione, cioè anche solo guardando il

Il mio volume di prelievo non è tutto il muscolo. Sento prevalentemente le sorgenti sotto gli elettrodi ma non certamente quelle più lontane e profonde. Sento un insieme di unità motorie limitato. La cute durante la contrazione, dato che il muscolo si deforma, scivola, quindi c'è una sorta di traslazione in una direzione o in un'altra delle fibre muscolari, mentre gli elettrodi sono attaccati alla cute, e le sorgenti si spostano rispetto agli elettrodi, quindi alcune sorgenti possono uscire dal volume di prelievo, altre vi possono entrare, mentre il volume di prelievo degli elettrodi è sempre lo stesso! Quindi a seconda di come cambia il muscolo geometricamente può essere che prima sento un certo pool di sorgenti e poi durante la contrazione questo cambia. Tornando al modello, se cambia il numero di unità motorie allora cambia P, se cambia P cambia la sommatoria, quindi cambia s, quindi non sarà WSS.

Quindi l'altra regola per avere un segnale EMG da considerare un processo stazionario almeno in senso lato è avere una contrazione a forza costante e isometrica, cioè col muscolo di dimensioni bloccate.

Questo spiega perché normalmente l'EMG in contrazioni volontarie non si consideri WSS a meno che non venga prelevato durante una contrazione isometrica a forza costante.

Per fare una contrazione isometrica basta vincolare l'articolazione in qualche modo, bloccandola in un opportuno sistema di misura, o dare al soggetto un compito motorio che non può vincere.

PSD di un segnale EMG

$$S(\omega, t, F) = R(\omega, d) \left[\sum_{i=1}^{P(F)} S_{u_i}(\omega, t) + \sum_{\substack{i,j=1 \\ i \neq j}}^{q(F)} S_{u_i, u_j}(\omega, t) \right]$$

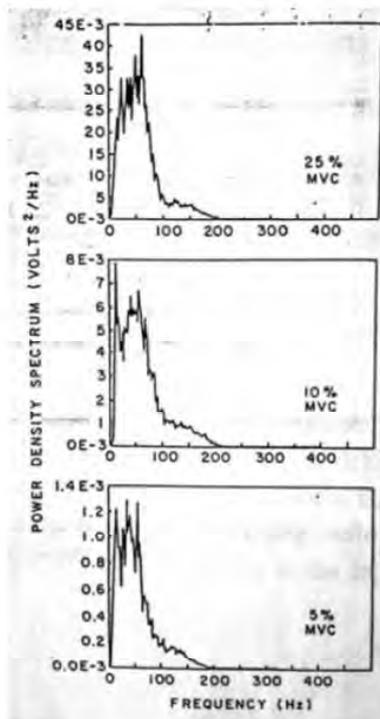
- d distanza interelettrodica
- P numero di u.m. di cui è costituito il segnale
- q numero di MUAPT correlati
- $R(\omega, d)$ funzione di trasferimento del sistema di prelievo

Osservare come il sistema di prelievo abbia grossa importanza e come compaia sia lo spettro dei MUAPT che il cross-spettro tra MUAPT differenti

È la somma degli spettri delle singole unità motoria. Se abbiamo un segnale fatto da una somma di segnali (ui sono i treni di potenziale d'azione delle singole unità motorie, quindi sono somme di segnali) se il segnale è scrivibile come somma di segnali, allora dato ha la FT è lineare, la FT di s non è altro che la somma delle trasformate di Fourier di u. Lo spettro però è equivalente al modulo quadro, quindi se misuriamo lo spettro scopriamo che corrisponde alla somma degli spettri quando i singoli segnali son scorrelati fra loro, mentre se correlati bisogna tenere conto della correlazione. La formula mi dice che lo spettro di s non è altro che la sommatoria degli spettri delle singole unità motorie, più gli eventuali spettri di unità motorie correlate, quindi q è il numero di unità motorie correlate fra loro, R la fdt del sistema di prelievo.

Unità motorie correlate vuol dire che il SNC può decidere di faar sì che alcune unità motorie lavorino in modo correlato, cioè se aumenta l'attività di un'unità motoria aumenta anche quella dell'altra. Dato che in una combinazione di segnali quando misuriamo lo spettro dobbiamo tener conto di eventuali correlazioni, a questo punto dobbiamo renderci conto che il SNC può indurre le unità motorie ad avere un'attività

ESEMPI DI PSD DEL SEGNALE D'INTERFERENZA



Grafici di un lavoro del 1976, MIT Cambridge, di N.J. Hogan. Vediamo tre densità spettrali di potenza del segnale EMG misurato dal bicipite omerale a diverse forze di contrazione: 25% MVC (=massima forza di contrazione volontaria) 10% MVC e 5% MVC. Queste contrazioni sono contrazioni a forza costante, quindi è stata misurata la massima forza di contrazione volontaria di un soggetto e poi gli si è detto (dandogli un feedback) di contrarre al 5%, al 10% e al 25% etc.

Come misuro la massima forza di contrazione? L'esercizio è stato fatto in questo modo: il braccio del soggetto era bloccato a un certo angolo articolare, due elettrodi posti sul bicipite omerale, gli si è fatta eseguire una contrazione per misurare la max forza di contrazione, il vincolo aveva un sensore di forza che misurava la forza articolare erogata dal soggetto. Negli sforzi successivi gli si è fatta fare una contrazione al 5% al 10% e al 25%, si è acquisito l'EMG e fatto il suo spettro.

L'articolazione è vincolata, l'avambraccio è vincolato a un piano, il manipolo che ho in mano misura la forza che sto esercitando. Non posso dire al soggetto contrai con tutta la forza che hai, perché c'è troppa variabilità e non posso misurarla una volta sola. Possiamo

aggiungere peso man mano e vedere finché non ce la fa più per trovare il carico, ma anche questo non è corretto perché non tiene conto dell'affaticamento.

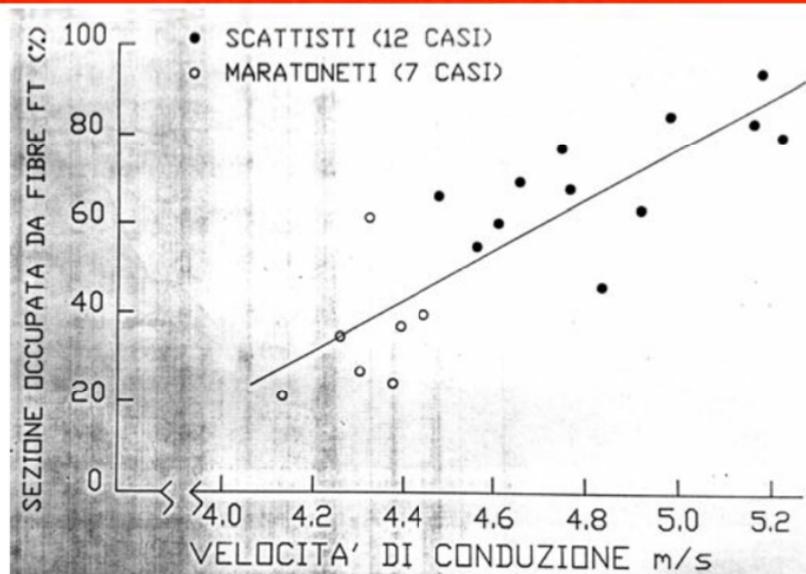
Quando si tratta di contrazioni volontarie, se faccio fare 3 contrazioni volontarie di seguito, c'è una certa variabilità rispetto alla forza che misuro nelle tre prove, che un po' è dovuta al fatto che il sistema si affatica e un po' al fatto che il SNC recluta in modo diverso e genera forze in modo diverso in base al nostro stato, es. se ci fanno in cazzare la forza è maggiore. Quindi definire qual è la massima forza di contrazione volontaria non è semplice, quell'MVC quindi viene misurato dalla media di tre contrazioni dette "motivazione", cioè in cui qualcuno dà una certa motivazione al soggetto.

NB: Il grafico ci dice che il segnale EMG non ha spettro piatto, ce lo aspettavamo, è approssimabile a un processo casuale non bianco, ma colorato. Il grosso della potenza è compresa tra 0 e 200 Hz, infatti l'EMG di superficie può arrivare al massimo a 500, ma se va bene a 350. La forma non varia al variare della forza di contrazione, ma varia il picco, infatti a forza di contrazione maggiore l'ampiezza dello spettro è maggiore, quindi l'area sottesa alla curva è aumenta all'aumentare della forza perché vengono reclutate più unità motorie oppure perché le unità motorie vengono sincronizzate tra loro. Quindi sappiamo che all'aumentare della forza di contrazione aumenta la potenza del segnale EMG, per uno dei due motivi, ma da quale dei due dipende dal soggetto, dal tipo di esercizio e da altre cose.

PROPRIETÀ DELL'EMG IN CONTRAZIONI VOLONTARIE

L'ampiezza va da pochi microvolt picco-picco a pochi millivolt picco-picco, dipende dal muscolo, muscoli piccoli hanno segnale EMG minore, quelli più estesi ce l'hanno di ampiezza maggiore. La banda va da pochissimi Hz fino a 500-1000 Hz (500 per prelievo di superficie e 1000 per prelievo invasivo) e se supponiamo che sia acquisito durante contrazioni volontarie isometriche e a forza costante può essere

Velocità di conduzione e fibre



L'unico parametro che possiamo usare per separarli se li vediamo è la massa muscolare. Ci sono delle variabili che differenziano l'uno dall'altro che si riconducono alla loro composizione e performance muscolare. In un grafico di dispersione in cui per ogni soggetto abbiamo la velocità di conduzione in ascisse e in ordinate la sezione muscolare (parametro parente del volume muscolare) vediamo che gli scattisti hanno velocità più alte, e facendo la retta di regressione (una misura di correlazione lineare) troviamo una buona correlazione tra velocità di conduzione e sezione muscolare. Quindi si scopre che chi è allenato a compiti motori che richiedono molta forza in un tempo molto breve (scattista) ha mediamente fibre muscolari più grandi e velocità di conduzione delle fibre muscolari molto grandi.

Quando parliamo di velocità di conduzione delle fibre muscolari, non parliamo di velocità alte come quelle nervose (40 m/s), in cui c'è la mielina, se togliamo la mielina nelle fibre fisiologiche la velocità è circa 4-4.5 m/s. In alcuni soggetti la velocità di conduzione è decisamente più alta e questo correla con l'aver fibre muscolari di dimensioni maggiori dette FT, cioè fast twitch (twich=velocità di contrazione), sono fibre molto veloci a contrarsi e rilassarsi quindi danno una forza molto grande in un tempo molto piccolo, ma quelle prestazioni durano per tempi molto piccoli. Tutti i nostri muscoli sono fatti da un mix di fibre lente e veloci, in proporzione diversa da un soggetto all'altro in base alla genetica e all'allenamento, ai compiti che si svolgono.

I maratoneti hanno invece fibre più lente e di diametro minore, ma molto resistenti alla fatica, quindi possono ripetere l'azione meccanica di contrazione e depolarizzazione senza incepparsi. Abbiamo descritto due tipi di fibre presenti nei muscoli:

- Quelle degli scattisti: fibre a metabolismo glicolitico
- Quelle dei maratoneti: fibre a metabolismo aerobico

Abbiamo tutti e due i tipi di fibre. La differenza è data dal fatto che le fibre aerobiche generano energia in presenza di O₂, hanno bisogno di O₂ per lavorare correttamente e scindere le sostanze nutritive. Le fibre a metabolismo glicolitico invece ossidano il glicogeno anche in assenza di ossigeno, ma senza ossigeno ovviamente non durano molto, quindi le prestazioni decadono.

Soglia aerobica: si fa l'analisi del gas espirato, quindi le concentrazioni di O₂ e CO₂ nell'espirato, funzione di come lavorano i muscoli. Anche a riposo ognuno di noi ha un suo consumo di ossigeno basale dato

17/04/2015

Essendo il muscolo fatto da fibre di tipo 1 o 2 o bianche e rosse (le prime sono a metabolismo glicolitico e sono più grosse e sono le più veloci e più potenti, le seconde quelle a metabolismo aerobico sono più lente, meno potenti ma più resistenti) la variabile fondamentale che comanda la capacità muscolare è la velocità di conduzione delle fibre muscolari: in elettromiografia velocità di conduzione è riferito alla velocità con cui il potenziale di depolarizzazione della fibra si propaga su essa.

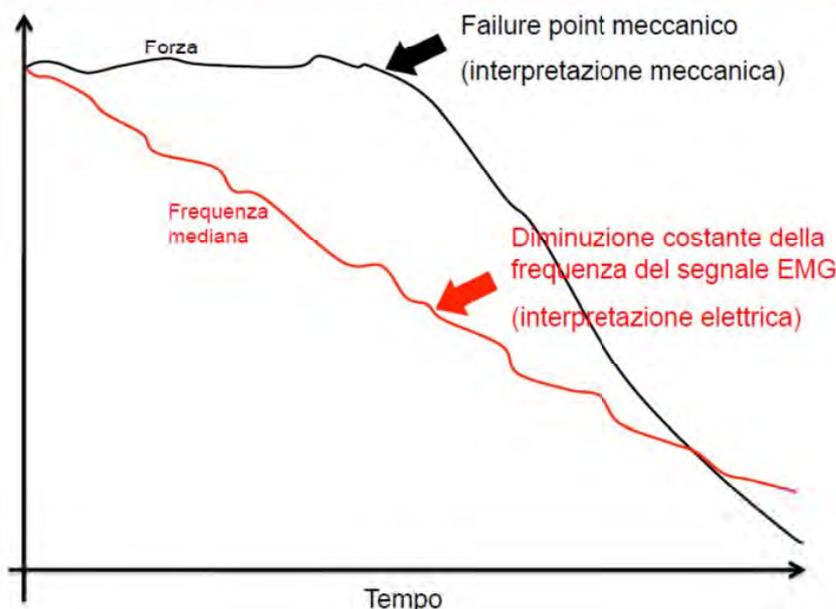
Correlazione tra sezione muscolare e velocità di conduzione tra gruppo di soggetti appartenenti a due popolazioni: compiti di minore intensità e lunga durata e viceversa. Per avere performance maggiori le fibre più grosse (bianche) hanno velocità di conduzione maggiore, e ciò si paga col fatto che tale fibra produce il massimo del suo lavoro per un tempo limitato, dopo di che si esaurisce, non riesce più a depolarizzarsi e contrarsi con le stesse prestazioni; al contrario le fibre rosse non sono in grado di erogare tanta potenza in un tempo molto breve, ma durano di più nel tempo. I nostri muscoli articolari sono una composizione di questi due tipi di fibre, che dipende in parte dall'allenamento, in parte dalla genetica.

La velocità di conduzione ha un taglio inferiore della intorno a 3.8-4.2 m/s; la velocità di conduzione è molto inferiore rispetto a quella delle fibre nervose, e questo è un valore limite, perché valori decisamente inferiori normalmente sono patologici, es. 3.5, 3.4, 3.2, non molto inferiori. Questo vuol dire che se vuoi dare una valutazione delle performance motoria, o una valutazione complessiva dello stato muscolare, certamente la velocità di conduzione è un parametro fondamentale che quindi bisogna saper misurare.

La velocità di conduzione delle fibre muscolari è direttamente coinvolta nel fenomeno della fatica muscolare localizzata: esiste una teoria che differenzia il concetto di fatica da quello classico di fatica che possiamo immaginare.

Grafico puramente qualitativo che mostra degli andamenti: grafico nero,

Fatica meccanica ed elettrofisiologica



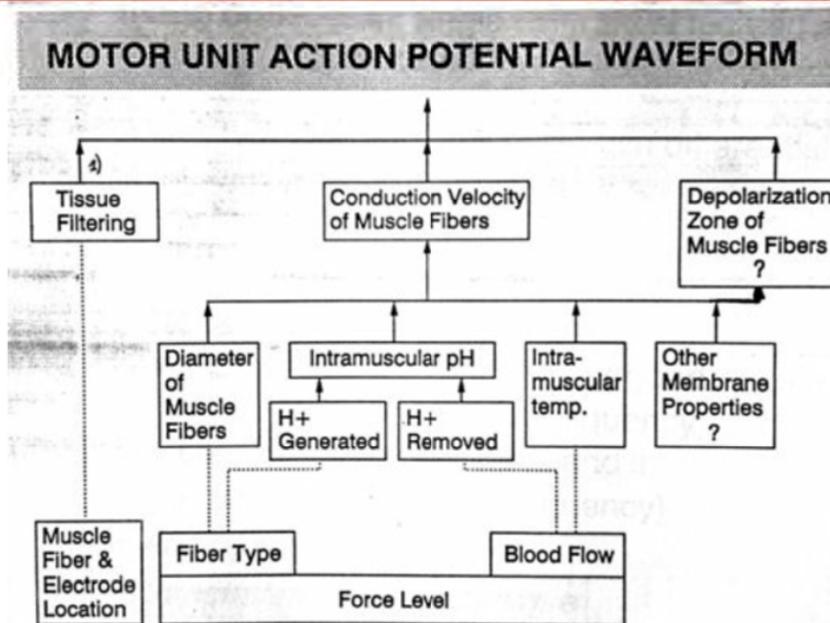
Supponiamo di avere un soggetto a cui faccio tenere un angolo di 90° e poi gli do un peso di 15 Kg, chiedendogli di non fare variare l'angolo articolare: per un po' di tempo il soggetto riesce, ma poi non riesce ad assolvere al compito e l'angolo inizia ad aumentare. Quindi il sistema per un po' di tempo riesce ad assolvere al compito motorio e poi dopo un po' c'è il failure point a partire dal quale le prestazioni

assone arriva alla giunzione neuromuscolare, ed è come se fosse una sinapsi: arriva un potenziale, fa qualcosa di chimico, riparte il potenziale sulla fibra, che crea la depolarizzazione. Dopo un po' il neurotrasmettitore si esaurisce o non è più sufficiente a creare la depolarizzazione. La giunzione non ha nulla a che vedere con la fisiologia del muscolo: quando parliamo di fatica muscolare localizzata parliamo di ciò che succede alle fibre e al muscolo in quanto costituito da fibre, invece la giunzione neuromuscolare sta a monte. A meno il compito non sia estremo o il soggetto patologico la giunzione neuromuscolare non si esaurisce, si affatica il muscolo per altre ragioni. Invece in alcune patologie il muscolo funziona ma si esaurisce la giunzione, e quindi il muscolo diventa disfunzionale.

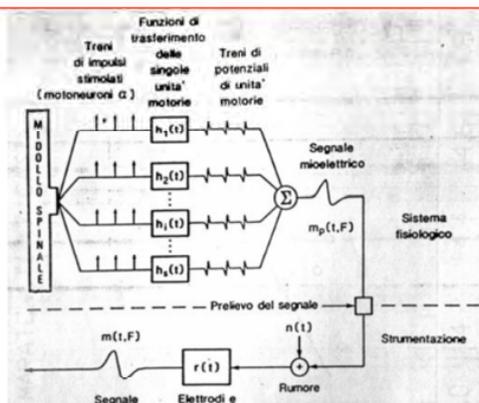
Il segnale EMG è relativo al blocco "fatica muscolare localizzata".

Inoltre dobbiamo considerare che quando pensiamo al fenomeno meccanico della fatica, cioè non riusciamo a vincere una resistenza, ad avere una forza etc quasi sempre dobbiamo considerare che quella variabile e quel compito motorio è il prodotto di un insieme di muscoli, non di un muscolo. Es. se misuro la forza con cui tiro un pallone, misuro la forza di una combinazione di muscoli, che vanno dall'anca al collo del piede che però riassumo con una sola variabile, la forza di impatto. Mentre con EMG io miro all'attività di un singolo muscolo, perché sono in grado di rilevare l'attività del muscolo che mi interessa. Quindi c'è anche un'implicazione pratica diversa tra le due interpretazioni del fenomeno di fatica.

Fattori agenti sulla forma del MUAP



Già da quando abbiamo fatti il modello del segnale EMG di superficie in contrazioni volontarie, potevamo intuire alcune cose, ci riferivamo a questo: unità motorie rappresentate con un sistema lineare con una

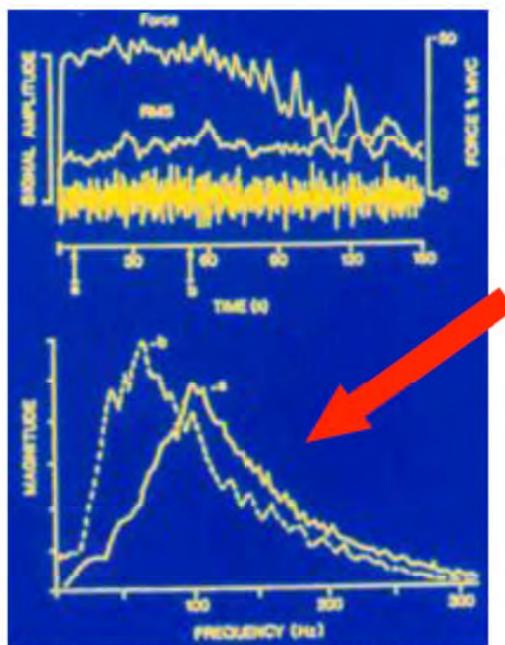


L'accumulo di ioni H⁺ ha vari effetti, tra cui ad esempio la generazione di un campo elettrico attorno alla fibra muscolare, questo rallenta il potenziale d'azione, che fa più fatica a propagarsi. Gli ioni H⁺ alterano anche un po' la conducibilità della fibra, quindi al netto la velocità di conduzione diminuisce. Quindi questo è il primo effetto fondamentale che ci aspettiamo da una contrazione muscolare prolungata se la consideriamo dal punto di vista elettrofisiologico, la diminuzione della velocità di conduzione.

Per soggetti come gli scattisti, con vdc di 5 m/s, le fibre lavorano in modo molto intenso e sono fibre bianche, producono molti prodotti di scarto e dopo pochi minuti le fibre bianche hanno una vdc molto ridotta e non sono più in grado di assolvere il compito meccanico efficacemente, quindi ad es. dopo 400 m lo scattista è morto. Invece chi è abituato a compiti aerobici (usa prevalentemente fibre rosse) produce meno H⁺ generano forza minore, c'è più flusso ematico e quindi il pH varia più lentamente, quindi la vdc decresce più lentamente, anche se partiamo da valori molto più bassi 4.2-4.3.

L'effetto fisiologico della fatica muscolare quindi è la diminuzione della velocità di conduzione.

All'instaurarsi di una contrazione ci sono da subito cambiamenti biochimici nel muscolo che causano una variazione della vdc e conseguentemente varia il segnale EMG: quindi se riesco a monitorare la vdc so dire qualcosa di come si affatica il muscolo dal punto di vista elettrico.



From J.V. Basmajian, C.J. De Luca, "Muscles Alive", Williams & Wilkins, 1985

Nel grafico c'è un soggetto che esegue una contrazione muscolare in modalità isometrica e a forza costante, partendo da circa 50%MCV (metà della sua massima forza di contrazione volontaria). Per un po' il soggetto assolve al compito dal pdv meccanico, poi vediamo a un certo punto il failure point meccanico, per cui la forza di contrazione inesorabilmente diminuisce. Sotto (grafico 1) c'è il segnale EMG misurato, per una contrazione di 150 s.

In due punti specifici, indicati da a = inizio contrazione e b= appena prima del failure point, è stato preso un pezzo di segnale EMG ed è stata calcolata la psd, riportata sotto (modulo della densità spettrale di potenza stimata in funzione della frequenza). Istante a =curva continua, istante b= tratteggiata.

MEDIANA: possiamo darla per variabili continue o discrete. È quel numero che divide la distribuzione in due parti aventi la stessa area, non mi interessa che ci sia lo stesso numero di campioni! Mi interessano i valori!!!

Quindi la frequenza mediana MDF divide lo spettro in due parti che hanno lo stessa area, per distribuzione continua quindi esiste una mediana. Per dati raggruppati la mediana non è univoca: la somma complessiva è ad es. 1+2+3+11+12, devo trovare il valore che la divide in due, e posso prendere il numero prima o dopo. Per distribuzione continua esiste la mediana, per dati raggruppati abbiamo l'incertezza di più o meno una classe. Vedi grafico istogramma.

Supponiamo che $P(f)$ (psd) sia continua, anche se in realtà in Matlab non lo è. Definiamo la mediana come il valore di frequenza per cui integrando dalla continua alla mediana misuro un'area che è la metà della totale:

$$\frac{1}{2} \int_0^{f_c/2} P(f)df = \int_0^{MDF} P(f)df$$

Definizione alternativa:

$$\int_0^{MDF} P(f)df = \int_{MDF}^{f_c/2} P(f)df$$

Dati P ed f calcolo la mediana come:

$A = \text{sum}(P)/2$ è l'area di riferimento, per cui per essere sicuri di essere arrivati alla mediana dobbiamo avere un'area pari ad A

s=0; i=1;

while s<=A

s=s+P(i);

i=i+1;

end

MDF=f(i);

se al ciclo i-esimo $s < A$ incremento di 1 e sommo ancora un campione, se al ciclo successivo è superiore avrò preso il valore dopo di mediana. Se voglio il campione precedente prendo $f(i-1)$. Attenzione perché questo

$RMS = \sqrt{fc \cdot \sum(x^2)/N}$ vale per qualsiasi segnale (segnale al quadrato, media, radice).

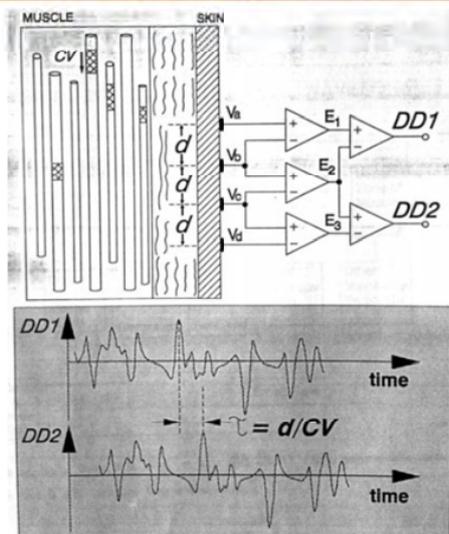
Ma per un segnale casuale (valor medio nullo) l'RMS coincide con la deviazione standard, perché l'RMS è la somma degli scarti quadratici rispetto al valor medio, che però è nullo, quindi per un segnale casuale è la std!

SOLO SE x è UN PROCESSO CASUALE $RMS=std(x)$!!!

Poi, la forza la misuriamo e poi calcoliamo la CV. Come?

Una velocità è spazio/tempo. La velocità di propagazione del potenziale sulle fibre per noi sarà la velocità che ci mette un potenziale a transitare sotto gli elettrodi. Normalmente quando si tratta di calcolare la vdc useremo la definizione fisica, spazio/tempo, tempo che ci mette un potenziale a percorrere quello spazio.

Prelievo di tipo doppio differenziale

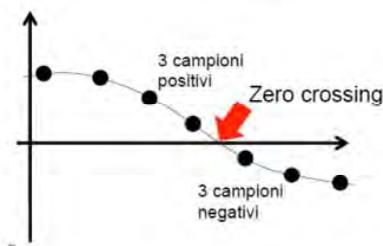
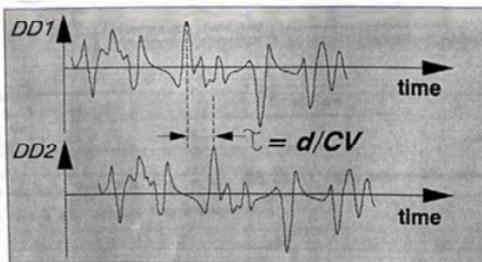


In figura i cilindri sono le fibre muscolari, quella segnata è la giunzione neuromuscolari e c'è un potenziale che corre lungo le fibre. Se faccio prelievo singolo differenziale ho un segnale EMG, e calcolare il tempo da lì è complicato perché gli elettrodi non sentono una fibra, quindi non vedo il potenziale che viaggia sotto gli elettrodi, ma vedo un insieme di potenziali che viaggiano contemporaneamente con velocità anche diverse magari. Quindi da un segnale singolo differenziale è difficile calcolare la velocità di conduzione.

Si preferisce usare un prelievo doppio differenziale, cioè anziché fare una volta la differenza tra i due elettrodi la facciamo due volte: il primo segnale doppio differenziale ($DD1$) è la differenza $E1-E2$, dove $E1$ è la differenza di potenziale tra l'elettrodo a e l'elettrodo b e $E2$ tra b e c. $DD2$ è la differenza $E2-E3$, dove $E3$ è la differenza tra gli elettrodi c e d. in totale per questo sistema doppio differenziale ho 4 elettrodi coinvolti. Il primo segnale $DD1$ è formato dai potenziali dei primi 3 elettrodi (a b c), $DD2$ dal potenziale degli ultimi tre (b c d). Supponendo che i 4 elettrodi siano stati messi equidistanti a distanza pari a d , l'insieme di elettrodi che genera $DD1$ rispetto a quello che genera $DD2$ è distante di una quantità pari a d (sono stati shiftati di una distanza pari a d).

Lo spazio quindi vale d , dobbiamo calcolare quanto tempo ci mette il segnale a propagarsi da un elettrodo all'altro, cioè per una distanza pari a d . Si scopre che ogni elettrodo ha un ritardo progressivo di

Zero Crossing technique



Tecnica che sfrutta poco l'informazione contenuta nel segnale.

Il rumore è un problema enorme per questa tecnica, poiché complica l'individuazione del punto di attraversamento dello zero.

La tecnica difficile da implementare, perché ad esempio in un segnale casuale è facile trovare in DD1 dove si ha l'attraversamento, ma devo poi beccare lo stesso punto sull'altro segnale, ma non so a priori di quanto sono ritardati i segnali. Visivamente è facile, perché vediamo tutto il segnale, ma un algoritmo non sa quali sono i punti omologhi tra i due segnali. Questa tecnica sfrutta molto poco le informazioni contenute nel segnale. È facile commettere errori, se si sbaglia il ritardo si sbaglia quindi la CV e il soggetto può passare erroneamente da normale a patologico.

Inoltre i segnali in figura sono segnali casuali con un buon SNR, ma se li sporchiamo con rumore, magari diverso tra DD1 e DD2, non vediamo più gli attraversamenti degli zeri. È una tecnica poco robusta a rumore sovrapposto. Si devono avere due segnali doppio differenziali con SNR stratosferico, 20 dB.

Alcuni dicono che non basta avere due campioni, uno positivo e uno negativo per vedere che c'è zero crossing, perché il segnale magari è variabile, ma ne servono almeno 3 consecutivi positivi e 3 consecutivi negativi (o 5 o 7 etc). Non la implementiamo perché porterebbe a stime sbagliate della CV.

Partendo dal presupposto che questa tecnica ha come difetto principale quello di non sfruttare l'informazione del segnale, dobbiamo usare altre tecniche che prendono in considerazione la struttura del segnale.

Ipotesi: DD2 è uguale a DD1, ma ritardato, cioè i campioni di DD2 sono gli stessi di DD1 ma compaiono dopo. Ci aspettiamo quindi che tra i due segnali ci sia una correlazione, perché sono lo stesso segnale, solo che uno è spostato sull'asse temporale. Misuriamo la correlazione mutua. Prendo DD1 e DD2 e calcolo la funzione di crosscorrelazione. La funzione di crosscorrelazione coincide con quella di autocorrelazione, ma il massimo è spostato in un valore sull'asse del ritardo corrispondente al ritardo tra i due segnali, infatti il campione di DD1 è massimamente correlato con quello di DD2 τ secondi dopo, se τ è il ritardo. Quindi mi aspetto che in questa ipotesi in cui DD1=DD2 questa funzione abbia un massimo assoluto in corrispondenza del ritardo (per un segnale con sé stesso la funzione di autocorrelazione ha un massimo nell'origine).

Come faccio a misurare CV? Prendo i segnali doppio differenziali, calcolo la funzione di crosscorrelazione, cerco il massimo e leggo il valore di ritardo in corrispondenza di quel massimo e quello è il valore di ritardo che devo usare per calcolare la CV.

La funzione di crosscorrelazione la calcoliamo nel discreto, imponendo un ritardo crescente man mano che trasliamo i segnali uno rispetto all'altro. Il minimo ritardo che possiamo imporre è 1 campione cioè in termini di tempo $1/f_c$ secondi, quindi la funzione di crosscorrelazione è fatta a punti che distano di una quantità temporale pari a $1/f_c$.

Una frequenza di campionamento che va bene per un segnale EMG di superficie è di circa 1KHz, perché la massima componente in frequenza è 500 Hz, e per Nyquist devo campionare almeno al doppio. Quindi in questa ipotesi ΔT tra i campioni è un 1 ms. Questo è un problema, perché devo misurare un ritardo di 2.5 ms e i campioni distano tra loro 1 ms, quindi se sbaglio anche solo di un campione a trovare il picco della funzione di crosscorrelazione sto sbagliando del 50 % a calcolare il ritardo e di conseguenza del 50% sulla stima della CV!!! Se sbaglio di un campione anziché $\tau=2.5$ ms misuro 3.5, oppure 1.5 ms, quindi nel primo caso otterrei una $CV=2.86$ m/s (sarei patologico!), nel secondo una $CV=6.6$ m/s (wowwww)! Quindi nel caso specifico del segnale EMG il problema diventa un problema di risoluzione: il segnale EMG nativamente è campionato a frequenza tale che la distanza in campioni del segnale si ripercuota sulla distanza nei campioni nella funzione di crosscorrelazione. L'asse temporale e l'asse del ritardo hanno la stessa frequenza di campionamento, quindi anche l'errore di un solo campione è drammatico e non me lo posso permettere!

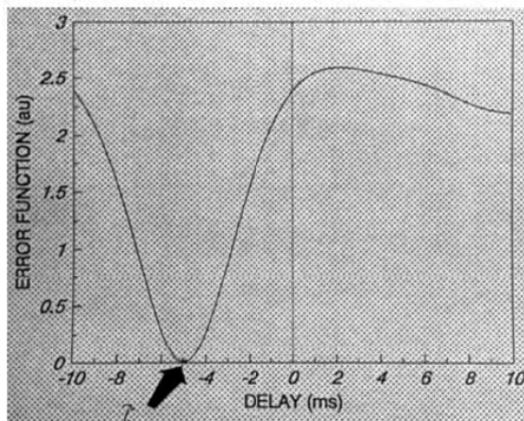
Per ottenere una buona risoluzione sul ritardo, 1ms non va bene. Se devo calcolare un ritardo di 2.5 ms voglio una risoluzione di 10 μ s che è la distanza temporale tra i campioni, cioè dovrei campionare a 100KHz. Non mi piace sovracampionare così tanto, perché a 1KHz ogni secondo produco 1000 campioni, a 100KHz 100000, e un tempo si aveva un problema di memoria, adesso questo non c'è molto, ma il vero problema è che sovracampionando se sull'asse delle frequenze la frequenza di Nyquist (metà della frequenza di campionamento) è lì, e il mio segnale ha una banda che finisce qui, in questo pezzo dell'asse delle frequenze che io sto campionando non c'è informazione, perché il segnale ha già terminato, ma non è che non c'è nulla, c'è rumore! Quindi in tutto quel pezzo che sto campionando non c'è segnale, ma c'è rumore, e quindi avrò un rapporto segnale-rumore di 10 dB, ovviamente, dato che gran parte della mia f_c campiona rumore, non posso avere un buon SNR!!!

Quindi per ovviare al problema della risoluzione dovrei sovracampionare, ma facendolo diminuisco l'SNR, torno ad avere il problema del rumore, che avevo sopra, quindi per correggere sotto ricado sopra, quindi dal punto di vista del segnale EMG funziona, ma non è il massimo.

Spectral Matching Technique (2)

Il valore che minimizza e^2 è esattamente pari al ritardo tra i due segnali.

Il minimo del funzionale e^2 teoricamente è nullo in corrispondenza del ritardo se i due segnali sono identici e semplicemente ritardati.



Come minimizzo ϕ ? Con un algoritmo iterativo noto come Metodo di Newton (vedi foglio).

Supponiamo di avere un certo $e^2(\phi)$, con minimo in $\phi=\theta$. Non possiamo calcolare tutti i valori di ϕ per trovare il minimo, ma si parte da un valore casuale ϕ' e calcolo $e^2(\phi')$ in corrispondenza di quel ϕ' , mettendo ϕ' nell'espressione. A questo punto calcoliamo la derivata, che in questo caso ha pendenza positiva, quindi sappiamo che il minimo è a sinistra rispetto al valore. Quindi so che ϕ' è più grande del minimo, quindi prendo l'intervallo $0-\phi'$ e lo divido in due, trovo il punto medio e lo chiamo ϕ'' , e calcolo il valore del funzionale in ϕ'' . Scopro che in questo punto il funzionale ha pendenza dall'altra parte, quindi il minimo è a destra, prendo l'intervallo tra ϕ'' e ϕ' e ne prendo la metà e la chiamo ϕ''' . Calcolo il valore di e^2 per ϕ''' , vedo che la tangente è positiva, quindi so che il minimo è a sinistra. Prendo l'intervallo, lo taglio a metà e lo chiamo ϕ'''' , etc finché non arrivo a convergenza quando la tangente è 0. Questo è il caso in cui la tecnica funziona bene.

Se la tecnica funziona male cado in un minimo locale, a seconda di come è fatto il funzionale es. figura

$x_1(t)+n_1(t)$ e $x_2(t)+n_2(t)$.

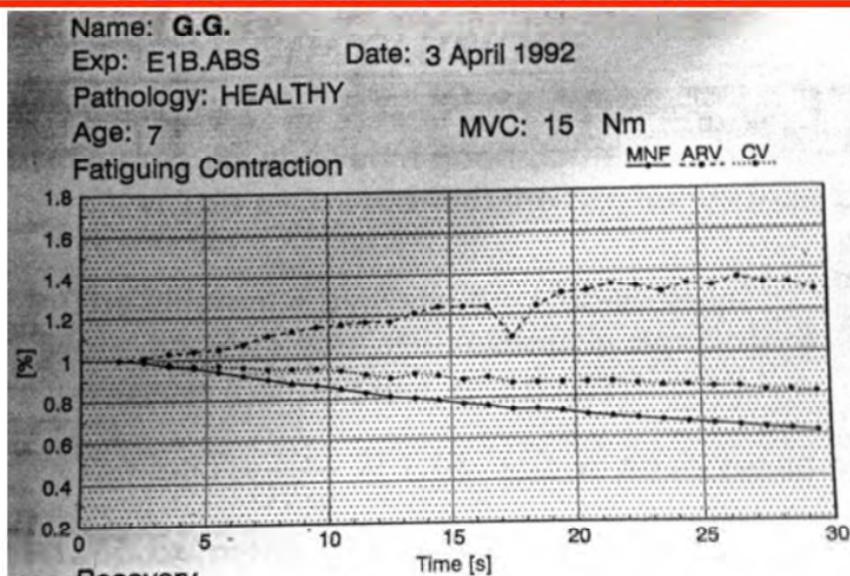
Nel tempo i rumori sono diversi, ma la loro trasformata di Fourier è la stessa se il rumore è gaussiano bianco, quindi si ha:

$$X_1 + N_1 = X_2 + N_2$$

Dove $N_1=N_2$. Quindi se i rumori che corrompono i due segnali dal punto di vista spettrale sono uguali, anche se ce n'è tanto e nel tempo sono diversi, se sono uguali in frequenza se ne vanno!

La robustezza della tecnica nasce proprio da questo, l'unico vero problema è che posso sbagliare a calcolare il minimo, quindi la questione è implementare la tecnica dello spectral matching completata da una serie di validazioni che faccio per capire se quel ritardo è coerente oppure no.

Fatigue Plot – Healthy subject



CV: 6.00 m/s ARV: 209 uV MNF: 101.4 Hz

LAB05: cosa succede al muscolo durante una contrazione prolungata nel tempo? Ci verrà dato un segnale EMG, singolo differenziale, di un soggetto che compie una contrazione isometrica a forza costante (siamo sicuri che il segnale è WSS). Poi ci vengono dati anche due segnali doppio differenziali DD1 e DD2) che servono per il calcolo della velocità di conduzione. La contrazione muscolare dura un certo numero di secondi. Dobbiamo rappresentare un grafico come questo, il fatigue plot, che mi dice cosa succede alla fatica muscolare durante l'esercizio. Si riportano la CV, un parametro spettrale (es. MNF in questo caso) e poi un parametro temporale (in questo caso ARV) dall'inizio alla fine della contrazione. Vedo come varia il segnale e quindi il muscolo durante un'contrazione che dura nel tempo.

Ho ad es. 30 s di segnale, i parametri non li calcolo su 30 s di segnale, ma lo spezzo a blocchi di 1 s e ogni punto del grafico è calcolato a un secondo di distanza.

Es. per calcolare la CV prendo i segnali doppio differenziali, calcolo il ritardo e calcolo CV, acchè il primo valore, lo faccio per ogni secondo, fino alla fine, in questo caso avrò 30 valori.

Se il muscolo si affatica la CV deve diminuire e quindi il fatigue plot viene normalizzato a 1, tutti i parametri vengono normalizzati. Es. se da 1 arrivo a 0.8 dico che la CV è diminuita del 20%, è semplicemente più facile

22/04/2015

Stima della velocità di conduzione delle fibre muscolari quando analizziamo il segnale EMG in un esperimento di fatica muscolare, o meglio quando vogliamo quantificare il fenomeno della fatica muscolare. Ci sono tecniche (si tratta di misurare il ritardo tra due segnali) che funzionano molto bene per misurare il ritardo tra due segnali, ma impongono delle condizioni sul segnale che però non sono rispettate dal segnale EMG.

Noi vogliamo imporre come unica condizione, quando parliamo di doppio differenziale, che i due segnali siano ragionevolmente simili, ritardati l'uno rispetto all'altro.

L'unica tecnica che misura la CV da un segnale singolo differenziale è la spectral dips, ma ha vari problemi, poi abbiamo visto la zero crossing in cui si calcola la distanza tra i due attraversamenti dello zero, ma è difficile da utilizzare, da automatizzare, non robusta dal punto di vista del rumore; poi quella basata sulla cross-correlazione che ha un significato fisico maggiore, ma ha il problema del rumore sovrapposto e inoltre l'accuratezza con cui misuriamo il ritardo è funzione della frequenza di campionamento.

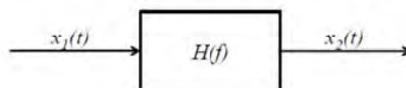
Spectral matching: ha il è quella che funziona meglio dal punto di vista del rumore, ma andando a cercare un minimo del funzionale d'errore, si incorre nel problema dei minimi relativi, quindi o si ripete più volte o si combina con la tecnica della cross-correlazione per vedere se il minimo valore di ritardo trovato è ragionevole.

Tecnica della Fase della funzione di trasferimento: tecnica non molto applicata, ma è interessante, perché ha un approccio modellistico che richiama quello dei filtri digitali e può essere utile. La spectral dips non è molto applicabile.

TECNICA DELLA FASE DELLA FUNZIONE DI TRASFERIMENTO, abbiamo la stessa ipotesi: due segnali x_1 e x_2 in cui x_2 è la replica ritardata di x_1 , sono ad esempio i classici doppio differenziali. Costruiamo un modello in cui supponiamo che x_2 sia l'uscita di un SLTI che ha come ingresso il segnale x_1 e in uscita ha x_2 che è uguale a x_1 ma ritardato. In questa ipotesi sappiamo che la fdt di questo SLTI è quel sistema che prende in ingresso il segnale e ti dà in uscita il segnale ritardato, cioè il ritardatore.

Phase of the Transfer Function Technique

Si supponga di avere due segnali $x_1(t)$ e $x_2(t) = x_1(t+\theta)$. Il segnale $x_2(t)$ si può pensare come l'uscita di un sistema LTI con funzione di trasferimento $H(f)$.



Se i due segnali di ingresso ed uscita sono identici e solamente ritardati uno rispetto all'altro, la funzione di trasferimento è un ritardatore puro.

$$\Phi(f) = \arctan \frac{\Im\{H(f)\}}{\Re\{H(f)\}} = af + b = \theta f + b$$

Il ritardatore discreto di k passi dal punto di vista dei segnali digitali è un filtro FIR, perché ha solo il numeratore ($\mathbf{H}(z)=z^{-k}$) quindi solo zeri. Il modulo di questa funzione di trasferimento è unitario, è un "passa tutto" dal pdv del modulo. La fase è lineare.

Vedi esercitazione sul segnale EEG: ci chiedevano di calcolare il ritardo che imponeva un certo filtro FIR, abbiamo visto che tale ritardo in campioni era proporzionale all'ordine del filtro stesso. Si può dimostrare che per un ritardatore discreto di k passi, cioè per una funzione di trasferimento

Se prendiamo il nostro EMG singolo differenziale e scriviamo il suo spettro, possiamo dire che lo spettro si può scrivere come prodotto di tre termini: lo spettro reale del segnale biologico (quello vero, che esce dalla sovrapposizione dei treni di unità motoria), a cui si sovrappone la fdt del tessuto, se presente, che è di tipo passa basso, a cui si sovrappone $R(f)$, cioè la fdt del sistema di prelievo. Si dimostra che per un sistema singolo differenziale, la $R(f)$ si può scrivere come $\sin \frac{\pi f d}{v}$ dove d è la distanza interelettroica, f è la frequenza e v è la velocità con cui i potenziali passano sotto gli elettrodi quando questi sono posti ad una certa distanza d .

Come sempre, quando abbiamo uno spettro che è il prodotto di tanti spettri, quando il termine $\sin \frac{\pi f d}{v}$ vale 0, vale 0 anche S , indipendentemente dal valore di G e di H , e dico che la fdt del sistema di prelievo a quell'andamento, cioè ha degli zeri che sono periodici, per come è fatta la funzione che è un seno. Quindi se $\sin \frac{\pi f d}{v}$ vale 0 sicuramente S vale 0, quindi uso una logica inversa per calcolare la CV: quindi guardo la psd di S , vedo quando c'è uno zero, quando c'è lo zero dico che si annulla il seno e allora a quel punto metto un argomento che annulla il seno stesso, a quel punto vado a calcolare la CV.

Vedi foglio

Supponiamo di avere questa psd casuale simil segnale EMG. Osservo che in questo punto c'è uno 0, e supponiamo che questo 0 sia ad una frequenza pari a f . Dico che quel primo 0 è causato dalla fdt dell'elettrodo. Quando è che $\sin \frac{\pi f d}{v}$ è pari a 0? Il seno vale 0 in 0, quindi ci sarà uno 0 lì, lo 0 successivo il seno ce l'ha per un argomento pari a π , quindi quando l'argomento è π quella roba è 0, quindi per un valore pari a f , quello che ho misurato, l'argomento del seno vale π , in modo che il seno valga 0 e $S(f)=0$.

$$\pi = \frac{\pi f' d}{v}$$

Dalla formula ricavo:

$$v = f' d$$

Cerco il primo zero su $S(f)$, so che lì l'argomento è π e lo eguaglio all'argomento del seno. Primo problema di questa tecnica: la tecnica non è tanto importante per il suo uso, ma è importante perché ci fa ragionare sui suoi vari problemi. Abbiamo detto che una distanza interelettroica ragionevole per EMG di superficie è intorno a **10mm**, perché se li mettiamo gli elettrodi troppo vicini sentiamo una porzione ridotta di muscolo, se troppo lontani sentiamo un volume di prelievo troppo grande. Se la **$d = 10 \text{ mm}$** e **$v = 4 \text{ m/s}$** (velocità media di un soggetto sano) scopriamo che **f è 400 Hz**, però siamo in una situazione del tipo:

piccole variazioni della velocità di conduzione. Distinguere 4 m/s da 3.9 m/s è molto importante per noi. Alcune di queste garantiscono risoluzione maggiore, altre ce l'hanno limitata.

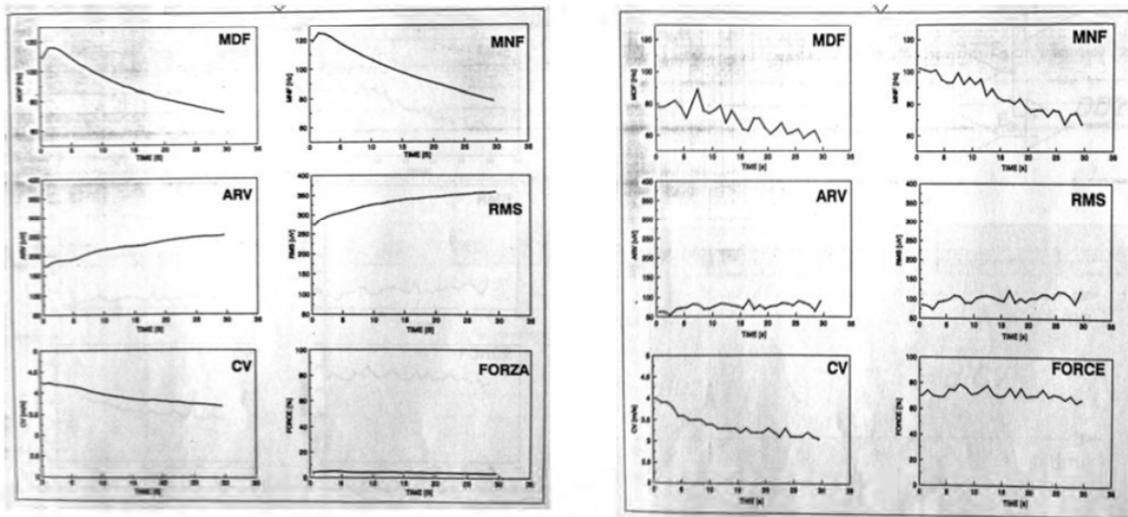
Alcune tecniche riescono a misurare una CV bidirezionale, altre no: noi abbiamo sempre detto che c'è un potenziale che si propaga e si propaga sotto gli elettrodi, io teoricamente non so in che verso, dipende da come sono orientati gli elettrodi, quindi vuol dire che eventualmente le tecniche devono riuscire a misurare anche un valore di ritardo negativo, quindi $x_1 = x_2(t + \theta)$ in cui θ può essere un valore negativo e quindi x_2 può essere in anticipo su x_1 ; alcune tecniche lo fanno, altre danno solo il modulo del ritardo, quindi non ti dicono se i potenziali vanno in media da una parte o dall'altra.

MANIFESTAZIONE DI FATICA MUSCOLARE

FATIGUE PLOT: comprende normalmente 3 parametri, uno è sempre la velocità di conduzione, il secondo è sempre un parametro temporale (ARV o RMS) e il terzo è un parametro spettrale (MNF o MDF). Noi facciamo l'ipotesi di contrazioni isometriche e a forza costante per garantire che il segnale sia ragionevolmente stazionario in senso lato.

La CV nel tempo diminuisce, è un dato di fatto. La MNF e la MDF dello spettro diminuiscono di conseguenza, perché quando la CV delle fibre diminuisce, la velocità di passaggio dei potenziali sotto gli elettrodi diminuisce, quindi il segnale EMG ha fronti sempre più lenti e quindi un contenuto in frequenza sempre più basso. Abbiamo visto nell'esperimento del 1975 che lo spettro si comprime verso le basse frequenze, quindi se misuro MDF o MNF di uno spettro che si sposta verso le basse frequenze, automaticamente durante la contrazione le due frequenze scendono.

ARV e RMS. L'andamento può dipendere dal tipo di muscolo, dal tipo di contrazione e dal tipo di "esercizio", nel senso che gli andamenti sono durante una contrazione stimolata elettricamente, se li guardo durante una contrazione volontaria gli andamenti sono leggermente a salire ma hanno una certa variabilità.



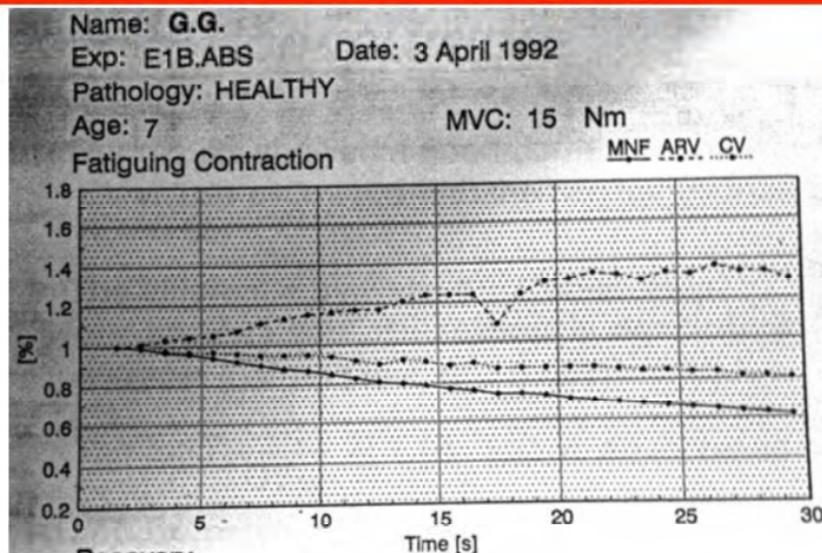
Perché ARV e RMS del segnale EMG tendono a salire durante una contrazione?

Il parametro temporale è il meno forte di tutti, la CV diminuisce, quindi si porta dietro il fatto che lo spettro si sposti verso le basse frequenze, MNF e MDF della psd diminuiscono e già qui possiamo vederlo bene se siamo bravi a fare stima spettrale e in base al rapporto SNR. Per quanto riguarda ARV e RMS che tendono a salire, non in tutti gli esercizi e non sempre, il segnale perde in frequenza aumenta di ampiezza, è questo che significa l'aumento di ARV e RMS! Come

l'esercitazione non ci viene un ARV o RMS crescente non è un problema, può succedere, come può succedere che per un po' aumenti e poi diminuisca, dipende da diversi parametri e diventa più complicato tener conto di tutto.

Esempio di fatigue plot:

Fatigue Plot – Healthy subject



CV: 6.00 m/s ARV: 209 uV MNF: 101.4 Hz

Per riuscire a calcolare in modo semplice come questi parametri variano durante la contrazione, il fatigue plot si normalizza all'inizio rispetto al valore iniziale che vale 1, e si vede come variano i parametri in funzione del tempo.

La CV diminuisce in modo quasi lineare al progredire della contrazione (che dura circa 30 s), la MNF diminuisce anche lei con andamento quasi lineare, un po' di più e l'ARV aumenta leggermente (a parte un punto in cui la stima non è proprio corretta). Questo è un caso da manuale di elettrofisiologia.

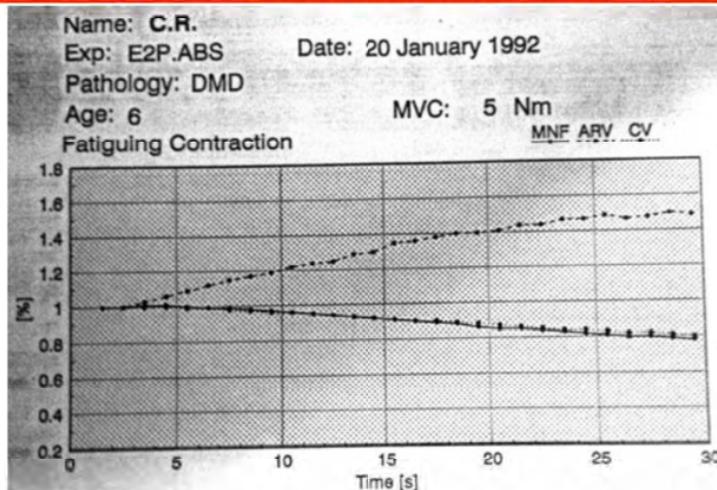
NB: Da questo grafico riesco a guardare bene l'andamento dei parametri, per capire se posso documentare un affaticamento muscolare, riesco misurare di quanto è variato ogni parametro, ma non ho più un riferimento rispetto ai valori iniziali, che quindi vengono riportati sotto, cioè sono i valori relativi al primo istante, quelli usati per denormalizzare tutto il resto. Da qui mi calcolo la CV al 30°s, so che è circa il 20% in meno del valore iniziale.

6 m/s l'abbiamo perché è un bambino di 7 anni, è difficile averla così alta. CV fisiologicamente è 3.8-4.3 m/s. I bambini hanno una composizione muscolare diversa e sono apparentemente infaticabili, ma si affaticano comunque. Quando un muscolo non si affatica durante un esercizio non è un buon segno.

MCV: massima coppia di contrazione volontaria è di 15 Nm, quindi probabilmente il muscolo è il tibiale anteriore, e abbiamo una flessione di caviglia. Supponiamo che sia questo esercizio, contrazione isometrica a forza costante al 50% MVC.

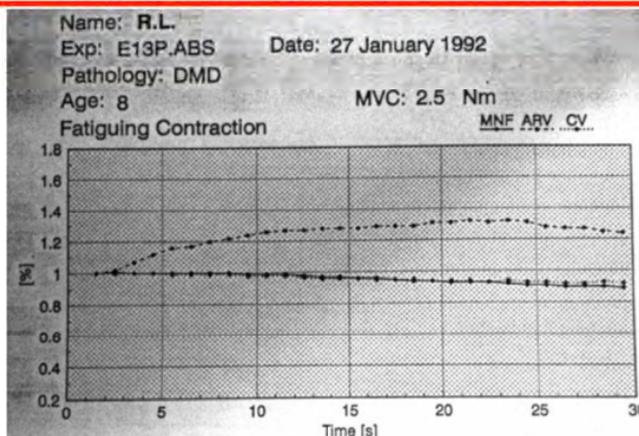
Questo grafico era di un soggetto sano.

Fatigue Plot – Duchenne Distrophy



CV: 2.96 m/s ARV: 214 uV MNF: 89.7 Hz

Fatigue Plot – Duchenne Distrophy



CV: 2.73 m/s ARV: 101 uV MNF: 88.5 Hz

Se prendiamo un caso in cui il livello di compromissione è maggiore (8 anni) da 5 siamo a 2.5 Nm, la coppia si è dimezzata, la CV non diminuisce più, partiamo da 2.7 m/s, la frequenza iniziale era 89 e adesso 88 Hz (più o meno la stessa) mentre adesso anche il parametro temporale ARV parte da valori molto bassi, cresce molto poco, cioè abbiamo sempre meno fibre muscolari attive con l'età. Quindi non è tanto il valore in sé, quanto l'andamento temporale, cioè quanto il muscolo risponde alla contrazione, è indicativo del grado di compromissione muscolare. Se un muscolo fatto affaticare si affatica di meno del 10% non va per nulla bene, quando lo stesso problema coinvolgerà un muscolo posturale il soggetto non cammina più, quando coinvolge un muscolo interno... l'aspettativa di vita per questa distrofia è molto scarsa.

Dato che la fatica muscolare può essere una misura di come reagisce il muscolo ad uno stimolo esterno, ad un esercizio che gli facciamo dare, ci dà il grado di compromissione del funzionamento muscolare, questo può essere un approccio utile per documentare l'invecchiamento. Un muscolo che invecchia è soggetto a sarcopenia, si parla ad esempio di sarcopenia primitiva dell'anziano, cioè il muscolo dell'anziano andando avanti con l'età va incontro a perdita di fibre muscolari. È una

24/04/2015

AUTOREGOLAZIONE CEREBRALE

Vedremo una serie di segnali un po' particolare, che pur avendo un contenuto spettrale importante, fanno un po' fatica ad essere analizzati con la stima spettrale tradizionale.

Cos'è l'autoregolazione cerebrale? Quando parliamo di autoregolazione intendiamo autoregolazione vascolare, quindi stiamo parlando di un fenomeno fisiologico che coinvolge un certo numero di parametri circolatori, o cmq legati alla circolazione e all'apporto di sostanze.

L'autoregolazione in linea di principio ce l'hanno tutti gli organi del corpo umano, sebbene alcuni organi abbiano un'autoregolazione molto spinta: è il caso del cervello che ha un'autoregolazione molto spinta.

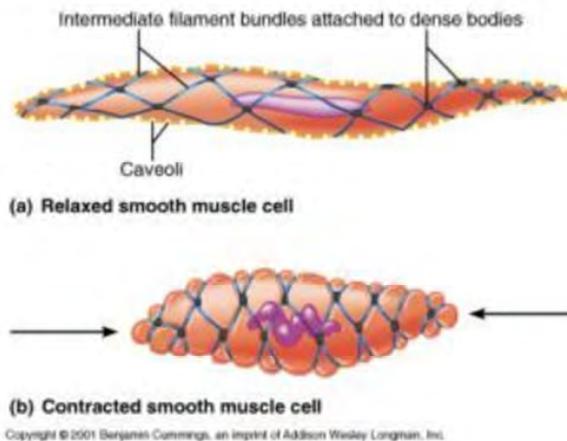
Abbiamo visto che il cervello si comporta in un modo un po' bizzarro, perché abbiamo visto che c'è una popolazione di neuroni che mette in atto la BARRIERA EMATOENCEFALICA, che fa sì che non tutto quello che circola può andare al cervello, ma solo cose ben particolari. Anche questa barriera ematica ce l'hanno tutti gli organi, però in alcuni organi è molto sfumata, in altri è molto importante. L'autoregolazione è molto importante per il cervello perché è fatto da cellule che patiscono particolarmente la mancanza di ossigeno e sostanze nutritive. La lezione scorsa abbiamo detto che il muscolo è perfettamente in grado di funzionare anche senza ossigeno, essendo fatto anche da fibre con metabolismo anaerobico, quindi se al muscolo gli togli ossigeno le fibre anaerobiche non se ne accorgono neanche e le fibre aerobiche dopo un po' vanno in crisi ma non muoiono e continuano a lavorare. I neuroni invece hanno solo meccanismo aerobico, e bastano pochissimi minuti in assenza di ossigeno perché il neurone degeneri permanentemente.

L'autoregolazione è quel fenomeno fisiologico che fa sì che l'apporto vascolare al cervello sia sempre costante, indipendentemente da quello che succede al resto del corpo. Se per qualche motivo si ha uno sbalzo di pressione, una riduzione dell'O₂, un aumento della CO₂... tutto quello che arriva al cervello deve rimanere il più possibile costante. Il motivo è che il cervello è fatto da cellule che, abbiamo visto, patiscono l'assenza, ma è anche privo di meccanismi che consentono l'immagazzinamento; il cervello non è in grado se gli diamo tanto O₂ o tanto glucosio di prenderlo, metterlo da parte e usarlo per quando non ne ha, cosa che fa benissimo il fegato, i reni, il muscolo... quindi il cervello ha bisogno di un apporto costante nel tempo, però i parametri circolatori costanti nel tempo non lo sono: al mattino abbiamo una pressione più bassa, la sera una pressione più alta, varia la temperatura, da quando siamo in piedi a quando ci sediamo cambia la pressione, quando andiamo a fare l'esame cambia la pressione.... Quindi tutto deve essere regolato in modo che l'apporto cerebrale rimanga costante. L'autoregolazione è quel fenomeno che fa sì che il flusso ematico cerebrale rimanga costante al variare della pressione: la pressione è la VARIABILE REGOLANTE, il flusso è la VARIABILE REGOLATA (cioè quella mantenuta costante al variare dell'altra).

Questo diagramma si chiama DIAGRAMMA DI AUTOREGOLAZIONE, e fa vedere come il flusso sanguigno cerebrale CBF (=cerebral blood flow) rimanga in un certo intervallo di pressione circa costante anche se la pressione ARTERIOSA varia. Esistono dei limiti intrinseci: se la pressione va sotto una certa soglia diviene insufficiente e non c'è la possibilità di autoregolare niente. Il limite inferiore dell'autoregolazione è circa 80 mmHg. Allo stesso modo se la pressione sale oltre il limite superiore dell'autoregolazione, il nostro sistema non è in grado di regolare la mandata ematica, è questo valore è circa 160 mmHg, inteso come pressione sistolica, perché la pressione diastolica è normalmente inferiore a 80 mmHg (oggi i cardiologi dicono che bisogna averne 75 mmHg).

Limiti di pressione:

organi o distretti con spiccate proprietà contrattili ci sono queste cellule muscolari lisce. Perché si chiamano lisce? Perché sono delle cellule contrattili con un corpo piuttosto allungato in condizioni di riposo e sono avvolti in una specie di rete proteica particolare (c'è chi dice che sono delle cellule in calze a rete!) e quando le cellule vanno in contrazione la maglia si stringe, e quindi diventano molto corte e cicciotte. Hanno una variazione della loro lunghezza molto importante, più importante di quella di altre cellule muscolari non lisce (generano una forte variazione del loro volume). Sono disseminate un po' dappertutto, ivi compreso nelle pareti vascolari di alcuni vasi, principalmente le arterie, ma anche le vene. Quando un vaso varia il suo calibro, quello che sta variando è lo stato di tensione delle cellule della muscolatura liscia. La vecchia battuta di qual è l'organo che varia macroscopicamente le sue dimensioni quando si contrae non fa ridere e cmq non è quello che si pensa: ovviamente è la pupilla che varia fino al 350% le sue dimensioni da tutto buio a tutto illuminato. Il gatto arriva a 550-600% (hanno una contrattilità ancora maggiore). Queste cellule sono le prime che degenerano in alcuni casi, e in alcuni organi non fa molto piacere (non in quelli che pensate voi!).



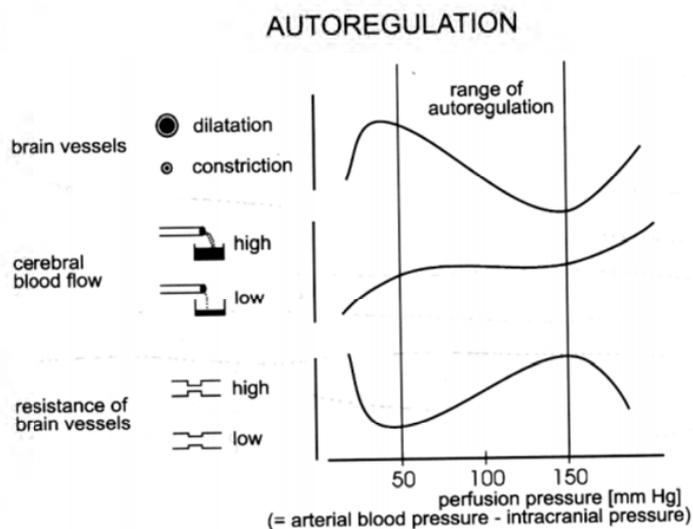
Si dice arteriosclerosi, aterosclerosi o arterosclerosi?

- Aterosclerosi: patologia che determina il deposito di lipidi all'interno delle pareti dei vasi e deriva dal greco ateros che vuol dire schiuma, perché l'aspetto del grasso quando si infiltra all'interno dei tessuti ha un aspetto schiumoso se visto al microscopio. Questo fa sì che il calibro del vaso si riduca
- Arteriosclerosi: modificazione della composizione della parete arteriosa e la perdita del tessuto muscolare che dà la proprietà di contrattilità. Le cellule della muscolatura liscia degenerano e vengono sostituite da tessuto fibroso che non ha grosse proprietà contrattili. Non è troppo diverso da quello che succede nella distrofia, ma non è un processo così pericoloso per la vita ed è molto lento. Le pareti diventano rigide.

Nel senso comune dire ad una persona che è un arteriosclerotico è un detto comune per dire che ha una patologia vascolare per cui dice robe sconclusionate perché non gli arriva la mandata corretta al cervello!

Come si fa a fare l'autoregolazione? Il grafico spiega cosa succede. Questa è la curva di autoregolazione vista prima. In ordinata c'è la pressione di perfusione, in ascissa ci sono riportate 3 cose diverse.

Al centro c'è il flusso ematico, che deve rimanere circa costante perlomeno all'interno di un certo intervallo di CPP.



radiazione, faccio finta che il resto dei tessuti non faccia niente, l'unico composto che assorbe la radiazione è l'emoglobina, e facendo il rapporto tra quanta luce mando e quanta me ne torna indietro tento di capire quanta emoglobina c'è. Questo si collega all'autoregolazione, che varia il flusso e quindi l'emoglobina. VANTAGGIO: La misura è totalmente NON INVASIVA. La radiazione IR non fanno niente, si usa pochissima potenza quindi non ci sono fenomeni di surriscaldamento, non sono radiazioni ionizzanti, per cui è un metodo comodo.

I pregi sono finiti qua, ora vedremo tutti i difetti. Questo dispositivo viene utilizzato perché dobbiamo ricordare che dentro il cervello c'è l'accoppiamento neurovascolare: fa sì che quando un neurone deve lavorare di più, perché deve svolgere una funzione, il sistema vascolare risponde immediatamente e gli fornisce quello che gli serve. Quindi volendo banalizzarlo, il neurone fa i capricci, ha bisogno di più risorse, il sistema vascolare gli fornisce subito la paghetta! Perché ci sia questo accoppiamento e come funziona è una cosa che non si sa, così come le cause per le quali questo accoppiamento si può perdere. È chiaro che la parte neuronale la vedo con l'EEG: l'EEG ha un contenuto armonico che è funzione dello stato del soggetto. La parte vascolare si può vedere con un certo numero di cose: si può fare una PET (ma i raggi gamma non piacciono a nessuno), una fMRI (non è un esame semplice e costa tanto). La Nirs è un metodo leggermente più veloce.

Come funziona? Quello che comanda tutto è la legge di Lambert-Beer, che è una legge esponenziale.

Facciamo un caso semplice: abbiamo un bicchiere contenente una soluzione al cui interno c'è qualcosa che assorbe la radiazione dell'IR. Supponiamo che di questa sostanza ce ne sia una concentrazione pari a c . nel nostro caso è l'emoglobina, che possiamo esprimere in mmol/l. da una parte del bicchiere iniettiamo radiazione elettromagnetica con intensità I_0 , poi essendoci le molecole del composto che assorbono le radiazioni a quella lunghezza d'onda, quando la radiazione esce dall'altra parte l'intensità sarà minore (I). Allora, se andiamo a calcolare l'attenuazione, che derivando dalla legge di Lambert-Beer è esponenziale, se lo vogliamo calcolare l'attenuazione esprimendolo come rapporto di I_0/I viene fuori il logaritmo, e scopriamo che l'attenuazione si può esprimere come prodotto di 3 termini

C =concentrazione sostanza

D =cammino ottico che i fotoni fanno nella soluzione (dimensione del provino)

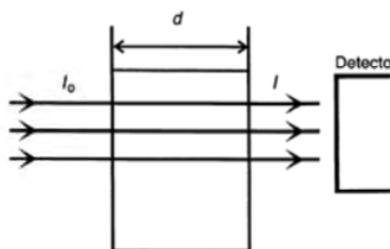
α =coeff di estinzione molare (dice una molecola di quel composto quanto assorbe in dB)

Soluzione di singolo cromoforo

$$A = \log \left[\frac{I_0}{I} \right] = \alpha \cdot c \cdot d = \mu_a \cdot d$$

α Coefficiente di estinzione

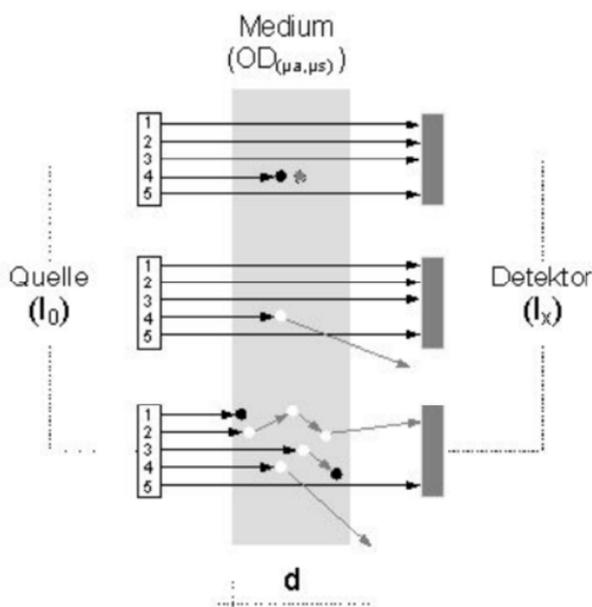
c Concentrazione del cromoforo



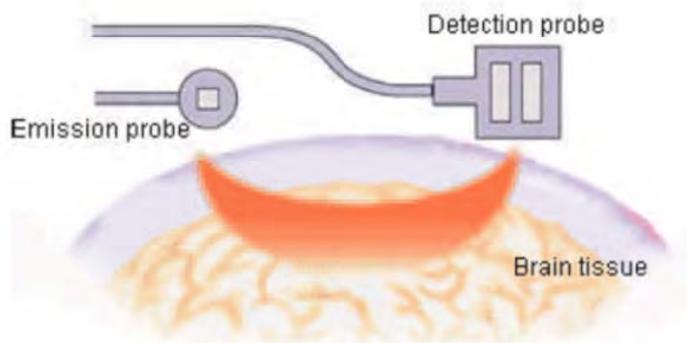
Il prodotto tra α e c è il coeff di attenuazione lineare o coeff di estinzione, simile a quello usato in radiologia e che si esprime in cm^{-1} .

Vediamo lo spettro di attenuazione dell' H_2O , o meglio coeff di attenuazione lineare dell'acqua. Nell'IR, fino a 900 l'acqua assorbe poco, poi ha un primo picco, quello piccolino, e se ci spostiamo ancora su aumenta ancora e quindi conviene restare sotto i 900 perché poi di acqua ce n'è talmente tanta che mangia tutto lei.

caso avrei una misura precisa dell'attenuazione se conosco quanti fotoni ho emesso). Succede che nel caso di scattering, alcuni fotoni deviano nel loro percorso e vanno in una zona in cui il ricevitore non c'è e quindi non li sento più e perdo questi fotoni. Nel caso reale ci sono sia fotoni assorbiti, sia fotoni scatterati più volte che arrivano cmq per caso al rivelatore, sia fotoni che vengono scatterati e poi assorbiti, fotoni che vengono scatterati e non arrivano al ricevitore e fotoni che non fanno un tubo e arrivano al ricevitore. L'80% dei fotoni viene scatterato dal tessuto cerebrale! Se metto una sorgente IR sulla cute, tutte le interazioni che ci sono tra i vari tessuti e liquidi determinano che l'80% dei fotoni vadano persi e questo comporta che non si possa fare trasmissione da una parte all'altra.



Succede che non posso quindi usare questa geometria, e allora ne si utilizza una diversa, che sfrutta proprio questa deflessione dei fotoni. Si utilizza un modello di questo tipo:



È stato dimostrato che se prendo una sorgente e la applico sulla cute, gran parte dei fotoni sterza, ma sterzano quasi tutti nella stessa direzione compiendo un percorso a banana chiamato proprio BANANA SHAPE, curvano, tornano indietro e quindi li vado a sentire a poca distanza dalla mia sorgente, a 3-4cm di distanza. Prendo così grossomodo l'80% dei fotoni, cosa che mi consente di limitare l'energia, ma

concentrazione, ma ci basta una variazione relativa. Allora si fa l'hp che nell'arco di tutto l'esperimento G rimanga costante, anche se ignota. Si chiama G perché è prevalentemente un fattore geometrico. Non ho idea di che valore abbia G, ma facendo l'ipotesi che resti costante nel tempo, l'unica cosa che varia nel tempo e che mi fa cambiare l'attenuazione che misuro è la variazione di concentrazione. Quindi si fa la differenza rispetto ad un istante iniziale preso come riferimento, per cui la legge diventa una legge differenziale e si ha:

$$\Delta A = \alpha \cdot \Delta c \cdot d \cdot B$$

La variazione di attenuazione che ho adesso rispetto ad un istante iniziale è uguale ad alfa*d*b*la variazione della concentrazione della sostanza che sto misurando.

Per cui la prima grossa limitazione è che con questa tecnica la quantità reale di Hb non la sapremo mai, ma sapremo quanto è cambiata rispetto ad un istante preso come riferimento. Misurazioni assolute non ne si riesce a fare, ma solo misurazioni relative.

Tutto sommato ci va bene, perché se vogliamo testare l'autoregolazione di un soggetto, lui è in condizioni normali, poi gli facciamo fare qualcosa, per cui lui se ha l'autoregolazione deve metterla in campo, e da quel momento cominciamo a valutare come variano le cose. Prendiamo un istante di riferimento e vediamo cosa succede più avanti rispetto a quell'istante. Questo per molti protocolli è un metodo che funziona.

Dobbiamo ricordarci che se mandiamo una radiazione ad una determinata lunghezza d'onda come 760 nm, non possiamo cmq dire che tutta l'attenuazione è dovuta alla desossiemoglobina, perché c'è ovviamente anche un coefficiente dovuto all'Hbo2, che è minore perché a questa lunghezza d'onda l'Hbo2 assorbe meno, però assorbe. Non possiamo dire che tutta l'attenuazione è dovuta alla co2, altrimenti facciamo i conti male. Stessa cosa se noi prendiamo una lunghezza d'onda di 900-950nm, siamo sul picco di assorbimento dell'emoglobina ossigenata, ma non è che l'Hbco2 non assorbe, anche lei ha un suo assorbimento. C'è una lunghezza d'onda poi in cui le due Hb assorbono entrambe la stessa cosa, per cui se misuro l'attenuazione, questa in parte è dovuta ad un Hb, in parte all'altra. se facciamo l'hp che ci siano sempre entrambe le forme di Hb dobbiamo tenerne conto e il sistema diviene così:

$$\left\{ \begin{array}{l} \Delta A_{\lambda_1} = \alpha_{\lambda_1}^{O_2Hb} \cdot \Delta c^{O_2Hb} \cdot d \cdot B + \alpha_{\lambda_1}^{HHb} \cdot \Delta c^{HHb} \cdot d \cdot B \\ \Delta A_{\lambda_2} = \alpha_{\lambda_2}^{O_2Hb} \cdot \Delta c^{O_2Hb} \cdot d \cdot B + \alpha_{\lambda_2}^{HHb} \cdot \Delta c^{HHb} \cdot d \cdot B \end{array} \right.$$

Vuol dire che l'attenuazione ad una lunghezza d'onda λ_1 è dovuto alla somma di due contributi: il primo è la legge di L-B che tiene conto del contributo dell'Hbo2, il secondo è il contributo che tiene conto dell'Hbco2, dove d e B sono uguali per entrambi, e gli alfa sono i coefficienti di assorbimento dell'Hbo2 a quella λ e dell'Hbco2 alla stessa λ . Però qui si hanno due incognite, cioè la concentrazione dell'o2 e la concentrazione della co2. Devo allora mandare necessariamente due λ diverse. Le due concentrazioni saranno le stesse indipendentemente dalla λ , quello che varia è ovviamente alfa. Adesso però ho 2 equazioni in 2 incognite e posso risolvere il sistema ricavando separatamente come varia la

allora i globuli rossi si accumulano nel reparto venoso e allora si accumula anche la CO_2 . Prima o poi la CO_2 quindi aumenta e se aumenta troppo il soggetto prima o poi ricomincia a respirare autonomamente e se non lo fa sviene. In altre parole suicidarsi trattenendo il respiro non è possibile, al massimo stramazzone. Questo è dovuto al fatto che l'organismo reagisce in due modi:

- O si rimpadronisce della respirazione facendoci svenire e così la respirazione diventa involontaria e ricominciamo a respirare
- Oppure ad un certo punto il diaframma ha delle contrazioni involontarie per cui dopo un po' arriva una botta diaframmatica che ci costringe ad inalare aria. È questo il motivo per cui si muore annegati: il colpo diaframmatico.

Quindi da questi segnali è facile verificare se il soggetto ha un'autoregolazione conservata oppure no.

Se invece che trattenere il respiro faccio apposta a respirare ad un ritmo forzato ponendomi in condizione di IPERVENTILAZIONE, questo è lo stimolo opposto che provoca l'effetto opposto, cioè la vasocostrizione e quindi a livello cerebrale l' O_2 diminuisce. Esempio classico: soggetto teso a cui gli si dice di fare un bel respiro. Questo si mette a fare respiri ampi e crolla così svenuto. Di ossigeno ce n'è troppo, il cervello non sa cosa farsene e chiude quindi la mandata vasocostringendo.

Il problema nasce dal fatto che se prendiamo questi segnali di variazione di concentrazione e andiamo a vedere come sono fatti nel dominio di Fourier scopre una cosa agghiacciante. La cosa che ci disturba non è tanto l'aspetto dello spettro ma i valori di frequenza! La banda P non centra niente con l'autoregolazione perché sono i picchi della pulsazione cardiaca. Quindi la banda di questo segnale arriva fino a 150-160 mHz. Volendo stare larghi 300mHz! Queste sono frequenze molto basse. Se ci danno un pezzo di questo segnale e ci dicono di fare stima spettrale la prima cosa che si valuta è ovviamente la risoluzione spettrale teorica. Se voglio vedere bene uno spettro che ha delle componenti che cominciano a 10-20 mHz e finiscono a 300mHz che risoluzione spettrale teorica voglio? 1mHz è ragionevole. Da 10 mHz a 100mHz con una risoluzione di 1mHz abbiamo 90 punti, non è che ci buttiamo via! Quanto segnale serve per avere una risoluzione spettrale teorica di 1 mHz? 1000s, quindi circa 15 minuti e un segnale biologico su 15 minuti non è stazionario. Questo è solo un esempio, ma di segnali biologici importanti che hanno una banda talmente stretta in frequenza, che se voglio vedere bene lo spettro diviso in bande ciascuna delle quali è correlata ad uno stato fisiologico del soggetto, per vederle bene devo avere una ris spettrale teorica talmente stretta che dovrei prendere un brano di segnale talmente largo che è impossibile che il mio segnale sia WSS. Quindi correlogramma e periodo gramma non sono utili per fare stima spettrale di questo tipo. Pian piano dobbiamo entrare nell'ottica che non esiste solo stima spettrale tradizionale, ma anche altri approcci in primis per venire in aiuto al caso di stima spettrale di segnali per i quali risoluzione e lunghezza non vanno d'accordo tra di loro.

precisamente quella energia. A noi quindi cosa conviene usare? Se guardiamo il sistema lineare ci conviene usare una sorgente laser, perché se ho un cromoforo e una lunghezza d'onda, perché α è un parametro imposto ad un valore preciso di lunghezza d'onda. Se ho un led che emette in uno spettro più ampio che α ci metto? Una media, ma non so bene qual è la campana di emissione del led e diviene molto complicato. Allora questa equazione viene risolta in un modo tanto più approssimato quanto più la sorgente è 'sporca', cioè quanto più non ha uno spettro concentrato in λ_1 e λ_2 . Quindi dal punto di vista matematico preferisco di gran lunga i diodi laser. Però i diodi laser si utilizzavano un tempo ed oggi si tenta di utilizzare dei banalissimi led. I motivi sono abbastanza ovvi: tra i diodi laser e i led c'è una differenza di prezzo di 2 ordini di grandezza! Quanto siamo disposti a pagare per un led che emetta una potenza nell'IR ragionevole con una banda non troppo larga? Spendiamo 10 euro, moltiplicato per 2 lunghezze d'onda o 3 a seconda di quanti cromofori si vogliono utilizzare, diciamo che ce la caviamo con 30 euro per la sorgente. Con 30 euro ci paghiamo solo il cavo fibra ottica del diodo laser, ma poi il diodo laser ci costa molto di più, perché per pilotare un led basta una resistenza e una pila, mentre per pilotare un diodo laser ci vuole un circuito di accensione e spegnimento, quindi diventa costoso, ci vuole la fibra ottica... ci sono tutta una serie di problematiche che ci portano a preferire quindi i led, ma c'è anche l'alternativa dei diodi laser. Il circuito di pilotaggio dipende da che sorgente si è scelto: banalissimo nel caso di led, un po' meno nel caso laser. Dall'altra parte, per il detector, le soluzioni sono quasi tutte univoche, cioè si utilizza un FOTODIODO o un FOTOTRANSISTOR (il circuito di utilizzo è pressappoco lo stesso). Si utilizza quindi un dispositivo che converte in corrente o tensione l'energia ottica che riceve in ingresso e tipicamente non è altro che un fotodiodo in polarizzazione inversa: il segnale in uscita dallo stadio di ricezione è proporzionale all'intensità ottica in ingresso.

Un fotodiodo distingue l'energia dei fotoni? Cioè sa se un fotone è stato emesso a 750nm o 800nm o 840nm?? No, un fotodiodo l'unica cosa che dice è che è arrivato un fotone, due fotoni, 10 fotoni, ma sicuramente non ha la risoluzione energetica sufficiente a dirmi a quale energia è arrivato quel fotone e quindi a quale lunghezza d'onda era. Però io in questo sistema devo misurare l'attenuazione ad una lunghezza d'onda λ_1 e ad una lunghezza d'onda λ_2 , come faccio? Se nella sorgente ci metto due λ (λ_1 e λ_2), due led o due laser non ci importa, entrambi mandano fotoni, l'Hb ossidata o ridotta si pappano parte di questi fotoni in base alla loro concentrazione, un certo numero di fotoni arriva al ricevitore, ma il ricevitore non distingue chi è chi. Per risolvere questo sistema ho bisogno però di misurare l'attenuazione alla λ_1 e l'attenuazione alla λ_2 , come faccio? Le due sorgenti le attivo in tempi diversi, e questo si chiama dispositivo TIME RESOLVED o A IMPULSI, il che vuol dire che in tempi diversi ho attivato sorgenti diverse. Tanto la distanza che devono coprire i fotoni non è tanto grande, di pochi cm, e i fotoni per definizione viaggiano alla velocità della luce cosicché quasi istantaneamente sul fotodiodo ho i fotoni che arrivano (con un ritardo che posso calcolare tanto conosco la velocità della luce). Quindi calcolo l'A a quella λ , dopodiché spengo la prima sorgente e accendo la seconda e calcolo l'A alla seconda λ . Quindi continuamente accendo e spengo e accendo e spengo. Considerando che devo continuamente accendere e spegnere conviene avere diodi laser o led? Idealmente se voglio fare le cose fatte bene a noi servirebbe che la sorgente emettesse un impulso sufficientemente energetico il più corto possibile. È evidente che più corto è l'impulso, maggiore è la f_c , perché più corto è l'impulso più velocemente riesco a ciclare tra λ_1 e λ_2 . Tra le due sorgenti quella che riesco ad accendere e spegnere più velocemente è ovviamente il laser. Costa una carota di soldi perché il circuito di pilotaggio arriva a sfiorare anche i 1000 euro, però se vogliamo possiamo ottenere anche degli impulsi dei picosecondi con tutta l'energia che voglio. Un led meno di 200-250ns è difficile: ha dei tempi molto più lunghi. Quindi i dispositivi time-resolved normalmente montano sorgenti laser per una questione tecnologico-costruttiva.

Questa però non è l'unica soluzione che ho a disposizione. Invece di accendere e spegnere le sorgenti che cmq mi costa energia e le sorgenti di qualunque tipo tutte le volte che vengono accese e spente e mandano un impulso abbastanza energetico non la prendono benissimo, per cui posso fare in modo che le sorgenti emettano di continuo. Ma come faccio allora a differenziare l'assorbimento alle diverse

sensoriale che man mano, con l'aumento dell'P media, cambiano il loro livello zero, il loro livello di base. Le variazioni veloci le leggono invece correttamente.

I barocettori misurano solo la P, ma per assicurare una buona autoregolazione non basta assicurare il flusso, perché questo è portatore di o₂ e sostanze nutritive, per cui bisogna leggerne la composizione.

- **CHEMOCETTORI:** misurano in realtà 3 cose. Misurano la P parziale di o₂ disciolto nel sangue, quella di co₂ e il pH ematico. In realtà basta misurare o₂ e co₂ e automaticamente ci si ricava il pH, perché l'o₂ è ossidante e la co₂ è riducente, per cui il pH non è altro che un bilanciamento tra le concentrazioni dei 2.

Si trovano più o meno nello stesso posto dei barocettori, comunicano con il nervo vago esattamente allo stesso modo, e quindi comunicano al SNC una variazione nella concentrazione dei 2 gas più importanti. Per cui si attivano quando l'o₂ scende sotto una certa soglia o quando la co₂ sale oltre una certa soglia e conseguentemente quando varia il pH. Il pH ematico è molto controllato: possiamo mangiare un quintale di arance fortemente ematiche ma il pH non cambia, al limite ci viene un'ulcera allo stomaco, ma il pH ematico non cambia.

Quindi con l'apnea volontaria i sensori leggono che la co₂ aumenta e l'o₂ diminuisce, lo comunicano al SNC, che vasodilata finché può, poi la co₂ continua a salire, l'o₂ a diminuire e il SNC va in condizioni di protezione imponendo una respirazione forzata o facendoci svenire.

È indubbio che l'osservazione dell'andamento temporale di questi segnali è indicativa di cosa succede dal punto di vista dell'autoregolazione. Però c'è un problema, perché il grafico dell'apnea della lezione scorsa è quello di un soggetto sano ed è un grafico 'da manuale': si vede che da quando il soggetto comincia l'apnea per un po' non succede niente, poi parte la vasodilatazione che fa aumentare l'ossigeno, poi non ce la fa più e quindi deve smettere di trattenere il respiro. Contestualmente la co₂ rimane ad un livello regolato fin tanto che comincia ad incrementare (per mancanza di o₂ e reflusso venoso) e deve ricominciare la respirazione. I grossi picchi che si vedono a valle della fine della respirazione, uno positivo e uno negativo sono dei picchi di reflusso, quindi c'è uno svuotamento del canale venoso perché il soggetto ricomincia a respirare e la co₂ va giù, c'è un aumento della P dato dal fatto che quando si riprende a respirare si ha una f cardiaca accelerata e quindi si ha un aumento di flusso istantaneo e poi si torna alle condizioni normali. Questo soggetto è stato scelto da manuale perché ha delle variazioni enormi delle concentrazioni a seguito degli stimoli. C'è però purtroppo una variabilità enorme da persona a persona perché ognuno ha un'autoregolazione un po' diversa e un suo funzionamento temporale un po' diverso. Inoltre stiamo guardando concentrazioni relative e non assolute. Quindi possiamo dire ad es che durante l'apnea la concentrazione di HbO₂ è variata di circa 3 umol/l. un altro soggetto potrebbe avere stessa dinamica ma con andamento un po' diverso. Di fatto le oscillazioni del segnale sono più o meno casuali, l'andamento complessivo no (l'o₂ deve aumentare e la co₂ deve aumentare solo alla fine) ma l'andamento temporale fa sì che il segnale sia prevalentemente a contenuto variabile e quindi prevalentemente un segnale casuale.

Per farne stima spettrale tradizionale per prima cosa bisogna togliere il valor medio. Secondo passo è togliere il trend del segnale se è presente. Questo è un segnale che ha delle oscillazioni, poi va su, ha delle altre oscillazioni e poi va giù, certo che è presente un **trend!** Allora quel trend alla stima spettrale fa male per tutti i motivi che ci siamo già detti: è una bassissima frequenza. Quindi normalmente i segnali NIRS in presenza di autoregolazione hanno sempre un trend che va tolto. Come lo vedremo quando guardiamo com'è fatto questo spettro.

Parte delle informazioni di questo segnale è contenuta nel dominio spettrale perché questo segnale è il prodotto del funzionamento di un certo numero di sensori che leggono come vanno le cose, ed è la

muscolatura liscia. Cioè la costante di t del fenomeno che regola la tensione dei vasi da quando i barocettori leggono il segnale a quando il SNC agisce sulle cellule della muscolatura per vasocontrarre o vasodilatare grossomodo la costante di t è proporzionale a 100mHz, e quindi è l'inverso di 100mHz. Questo è quindi la frequenza propria di questo sistema di controllo. Altri però dicono che non è vero e che è stato dimostrato che nel SNC c'è un oscillatore che funziona autonomamente e che con una f di 100mHz regola le cose con autonomia. Sia quello che sia ma all'interno di questa banda c'è l'impronta del sistema nervoso simpatico. State certi che se qualcuno ha un problema nervoso o vascolare la potenza nella banda m cambia rispetto ad un soggetto sano, qualunque sia lo stimolo (apnea, esercizio...).

Siccome poi questo è un segnale di concentrazione di HbO_2 e di $HbCO_2$, la percentuale di questi due tipi di Hb cambia in base alla respirazione, che è forse il più grosso stimolo. Allora siamo ad un valore di f che va da 150 mHz fino a 200-300mHz che sono un po' le frequenze tipiche degli atti respiratori di un soggetto. La f cardiaca di un soggetto sano a riposo è di circa 1 battito al s e cioè 1 Hz (60 battiti al minuto). Ma quanti atti respiratori si fanno al minuto? 15 atti respiratori al minuto, quindi 1 respiro ogni 4 secondi e quindi una f di 250 mHz. Questo è il fenomeno della respirazione, che oltre a variare l'apporto di O_2 e CO_2 e quindi influire sui chemocettori, varia la pressione intratoracica e con essa si ha una variazione della stasi o del reflusso venoso. Quindi anche questa modula la concentrazione che andiamo a leggere.

Quando parliamo di autoregolazione, ovviamente le bande principali sono la b e la m perché la respirazione siamo già bravissimi a misurarla in un altro modo e non c'è bisogno di fare tutto sto casino per vedere l'effetto della respirazione. Non siamo particolarmente interessati alla banda R , anche se c'è, e siamo ancor meno interessati alla banda P che è quella della pulsatilità cardiaca, che se vogliamo studiarla facciamo un ECG.

Quindi questo segnale ha diverse info nel dominio del t e nel dominio della f , ma per l'autoregolazione quello che ci interessa nel dom della f è compreso tra 10 mHz a 140-150mHz e quindi abbiamo un segnale con un contenuto frequenziale molto basso di cui ci interessano le f ancora più basse! Quindi al volta scorsa abbiamo detto che per uno spettro di questo tipo la volta scorsa si era detto che si voleva una risoluzione spettrale di 10 mHz, ma in realtà al minimo 10 mHz, forse anche meno! Il che significa avere dei monitoraggi che durano più di 10 minuti. Su 10-30 minuti non possiamo, per com'è fatto il sistema fisiologico che lo produce, non possiamo fare l'ipotesi che il segnale sia WSS. Sembra che si sia senza via di uscita, perché Fourier ci impone un rapporto tra la dimensione temporale del segnale e la risoluzione in frequenza, noi ci imponiamo una risoluzione in frequenza minima altrimenti non vediamo bene quello spettro lì, però se utilizziamo quella risoluzione violiamo W-K e se accorciamo il segnale per non violare W-K non abbiamo la risoluzione che ci serve per cui la stima spettrale è totalmente inutile.

Da qua non se ne esce con la stima spettrale tradizionale. Ci sono quindi dei casi in cui la stima spettrale tradizionale è insufficiente per i nostri scopi, e nell'ambito dei segnali biomedici capita molto più spesso che in altri ambiti: è molto difficile che in settori come le telecomunicazioni, la geologia, l'astrofisica ci siano problematiche di questo tipo, perché si ha a che fare con sistemi che rimangono stazionari per tempi molto lunghi, ma i sistemi fisiologici impongono una variabilità molto elevata! Questo è il motivo per cui accanto alla stima spettrale tradizionale si sono evolute altre forme di stima spettrale che ora introdurremo.

In ambito clinico un tempo questo problema era chiaro, ma non avendo a disposizione strumenti di stima spettrale alternativa al periodogramma e correlogramma che lavorassero sufficientemente bene e che fossero facili da implementare si tentava di stazionarizzare il sistema fisiologico. Abbiamo visto che l'autoregolazione dovuta all'apnea ha delle variazioni quasi istantanee, perché le concentrazioni variano velocemente a seguito della vasocostrizione o della vasodilatazione. Un tempo, invece che utilizzare l'apnea che è uno stimolo non stazionario, si utilizzava una vasodilatazione o una

Fine lezione 27/04/2015 (da minuto 1:00:53)

STIMA SPETTRALE PARAMETRICA

- **GENERALITA' SULLA STIMA PARAMETRICA**

SLIDE 4: STIMA SPETTRALE TRADIZIONALE: POSSIBILI PROBLEMATICHE

Fermo restando che questa cosa per noi non è rilevante, nel senso che si può fare come ci pare, ma la cosa più semplice (e per certi aspetti è anche più corretta) è quella di andare a misurare il contenuto spettrale del segnale che ci interessa, nelle condizioni reali del sistema fisiologico. Quindi non si va ad alterare il sistema fisiologico (che non ha senso), ma bisogna dotarsi di sistemi di analisi che vadano bene per il segnale che ci interessa.

E allora bisogna introdurre un nuovo aspetto teorico, che è il gemello della stima spettrale tradizionale => noi fino ad adesso abbiamo visto questa, e abbiamo detto che è anche stima spettrale non parametrica, perché si differenzia proprio da quest'altra stima spettrale parametrica e che serve per effettuare l'analisi spettrale in condizioni un po' particolari, cioè quando l'analisi spettrale basata su periodogramma e correlogramma non fornisce informazioni sufficienti, cioè non ci da prestazioni ragionevoli.

La stima spettrale parametrica si usa tutte le volte in cui la stima spettrale tradizionale funziona male => nessuno è così scemo da usare la stima parametrica quando funziona bene il periodogramma o il correlogramma, perché è una complicazione, ma si fa quando non se ne può fare a meno.

Un caso è risolvere il problema a cui eravamo arrivati prima, cioè necessità di una certa risoluzione, ma impossibilità pratica di ottenerla senza violare il teorema di W-K. Quella certamente è una situazione che ci indica di usare stima spettrale parametrica. Quindi: 1° caso quando si ha un problema di risoluzione, ovvero quando il segnale è eccessivamente corto oppure ha delle caratteristiche per cui è incompatibile con la risoluzione spettrale teorica necessaria.

Il 2° caso, che però è un caso molto minore a livello di importanza (quello di risoluzione è il caso fondamentale: tutte le volte in cui siamo nel 1° caso, non si può fare a meno di usare questa stima parametrica) => ci possono essere segnali con frequenza di campionamento non costante. Noi fino ad adesso abbiamo visto il segnale ECG e lo abbiamo campionato a 250 Hz, 300 Hz, 500 Hz; il segnale EMG lo abbiamo campionato a 1 kHz; il segnale EEG lo abbiamo campionato a 500 Hz... => questo vuol dire che tutti i campioni nel tempo sono equidistanti e hanno una distanza temporale che è l'inverso della frequenza di campionamento. Così è comodo, ma ci sono dei segnali fisiologici che questa caratteristica non la consentono, cioè ci sono dei segnali fisiologici dove i campioni, nel dominio del tempo, non hanno una distanza temporale costante perché vengono quando vengono.

Esempio molto semplice: supponiamo di misurare per un protocollo di medicina dello sport, la concentrazione di anidride carbonica nell'espriato di un soggetto che sta correndo sul tapis roulant (in condizioni controllate). Il soggetto in questo caso espira quando gli pare a lui, non espira a un ritmo preciso: in alcune fasi avrà un ritmo più lento, in altre più veloce e comunque un respiro dista dall'altro, una quantità variabile di tempo.

Altro esempio: voglio misurare la velocità ematica all'interno dell'arco aortico, a ogni onda sistolica. Ogni onda sistolica si ha a ogni contrazione cardiaca, ma il cuore non si contra a un ritmo perfettamente costante, cioè ogni battito cardiaco non è esattamente equidistante temporalmente l'uno dall'altro e la frequenza cardiaca può variare.

Allora ci sono dei segnali fisiologici che non hanno una frequenza di campionamento costante nel tempo. Il teorema di Nyquist inoltre non dice che la frequenza di campionamento deve essere costante, ma che deve essere superiore due volte della banda del segnale, non che sia costante. Quindi un segnale discreto

come cercare di farlo in modo corretto. Poi si vedrà che ci sono dei limiti su cui non si può fare niente, però ci sono dei casi in cui si può fare per alcuni segnali, anche abbastanza complicati, una stima spettrale anche abbastanza ragionevole.

29/04/2015

riepilogo...

Abbiamo introdotto quella che è la stima spettrale parametrica e abbiamo detto un paio di casi in cui la stima spettrale tradizionale non funziona particolarmente bene: il vero caso limitante è quando il segnale ha delle caratteristiche tali per cui con la stima spettrale tradizionale, essendo vincolati al fatto che la risoluzione spettrale è l'inverso della lunghezza temporale del segnale, non riusciamo ad avere degli spettri di buona qualità, o perché abbiamo bisogno di una risoluzione molto spinta o perché il segnale è corto: ci sono dei segnali fisiologici che hanno una durata abbastanza breve (tipo alcuni potenziali) e quindi riuscire a fare una stima spettrale ragionevole, è complicato. Altri casi sono quelli che possono avere una frequenza di campionamento non costante, variabile perché si campiona ogni volta che avviene un effetto biologico che non è controllabile. Allora in questi casi si può ovviare alle limitazioni introdotte dalla stima spettrale tradizionale utilizzando un approccio di tipo parametrico. È stata già fatta una precisazione importante: l'approccio parametrico lo usiamo sempre e solo quando quello tradizionale funziona male => la stima spettrale parametrica non è mai la prima soluzione, è sempre la seconda. Come seconda cosa si è detto che è coinvolto un approccio di tipo modellistico, cioè la stima spettrale parametrica di fatto vedremo che vuol dire calcolare un certo numero di parametri di un modello, che noi usiamo per rappresentare il segnale. Allora come sempre succede, i modelli funzionano benissimo quando sono semplici e si riesce a controllarli, ma quando sono modelli complicati o non riesco a controllarli, il loro funzionamento diventa un po' più difficile. Quindi si vedrà che va tutto bene quando ci sono segnali facilmente modellizzabili, mentre quando ci sono segnali non facilmente modellizzabili, anche la stima parametrica diventa complicata.

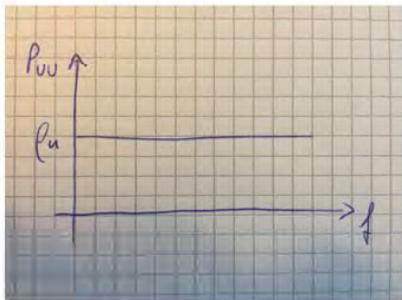
La terza cosa che non abbiamo detto la scorsa volta è la seguente: nelle slide è tutta matematica o poco più, tutte formule... => però non è importante ricordarsi a memoria le formule che ci sono scritte (anche se sono equazioni lineari e la notazione è semplice), ma il concetto fondamentale è capire a cosa servono, perché poi quando concretamente si farà stima spettrale parametrica, a quelle formule di passa sopra..però se non si cosa c'è sotto, non faremo una buona stima spettrale.

Questa è una cosa su cui bisogna fare attenzione, perché se ci sono un certo numero di criteri o di osservazioni che si possono fare per stabilire se un periodogramma o un correlogramma è di buona qualità o meno (cioè se abbiamo ottenuto una stima ragionevole o no), per quanto riguarda la stima spettrale parametrica è un pochino più difficile => va bene non fare tanto i fiscali sulla matematica, ma se poi si usano le istruzioni Matlab come scimmie che pigiano solo i tasti, non funziona tanto bene perché è facile ottenere un risultato che quasi mai è quello giusto (come sempre accade...). Quindi noi descriveremo quali sono gli aspetti matematici che vedremo, però i dettagli matematici (anche se non verranno chiesti all'esame) bisogna conoscerli perché sono quelli che guidano le scelte che si fanno dal punto di vista pratico.

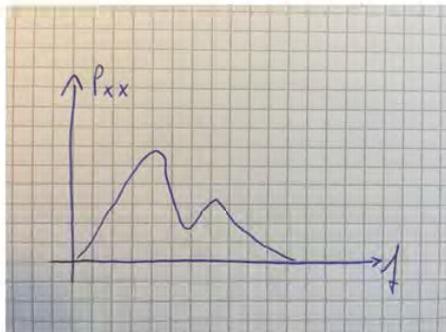
Slide 5: MODELLIZZAZIONE DI UN PROCESSO CASUALE

Si è detto che c'è di mezzo un processo di modellizzazione che si articola in 4 fasi:

- 1) SCELTA DI UN MODELLO APPROPRIATO (AR, MA, ARMA)
- 2) SCELTA DELL'ORDINE DEL MODELLO
- 3) STIMA DEI PARAMETRI DEL MODELLO
- 4) CALCOLO DELLA PSD A PARTIRE DAL MODELLO



In uscita invece si ha che la PSD del nostro segnale ha l'aspetto che ha:



=> Questo modello, se lo vediamo dal punto di vista "spettrale", cioè vediamo l'effetto che fa in frequenza, ci dice che: parto da uno spettro che è una costante (è piatto, non ha forma) in ingresso e ottengo in uscita qualcosa che ha forma e questa forma gliel'avrà data logicamente la funzione di trasferimento del sistema lineare, che imponendo le sue radici ha imposto una forma allo spettro medesimo del segnale. Quindi il responsabile della forma che ha la PSD del segnale in uscita, è la funzione di trasferimento del modello lineare. Il concetto è che $x[n]$ è un segnale scomodo, perché è corto, perché ha uno spettro molto basso, perché non riesce a mettere d'accordo la risoluzione con la durata temporale...allora $x[n]$ lo uso solo per costruire il modello e poi si sa già che la forma dello spettro è dipendente da $H(z)$ => tanto vale che una volta fatto il modello, lavoro su $H(z)$ e ignoro il segnale. Quindi $x[n]$ serve solo per costruire il modello lineare e poi verrà dimenticato e la stima della PSD deriva da $H(z)$.

Questo ragionamento, che è ovvio per un processo casuale gaussiano bianco in ingresso perché ha uno spettro piatto, si può dimostrare matematicamente per qualunque processo casuale; cioè si può dimostrare che dato uno spettro all'ingresso, posizionando opportunamente poli e zeri, riesco a modificare lo spettro che ho in uscita. Allora il succo del discorso è quello di riuscire a realizzare un modello ragionevole del mio segnale, in modo che questo modello possa poi essere usato per dare una stima dello spettro del segnale medesimo.

Allora facciamo un'ipotesi: supponiamo che il modello LTI abbia in realtà una funzione di trasferimento $H(z)$ che possiamo scrivere come un modello autoregressivo a media mobile (ARMA) di ordine rispettivamente p e q (p è l'ordine della parte AR, q della parte MA) => se volessi scrivere l'uscita $x[n]$ in funzione dell'ingresso $u[n]$, dato che la risposta all'impulso è la risposta all'impulso classica di un modello AR o MA, la relazione è:

$$\begin{aligned}
 x[n] &= -\sum_{k=1}^p a[k]x[n-k] + \sum_{k=0}^q b[k]u[n-k] \\
 &= \sum_{k=0}^{\infty} h[k]u[n-k]
 \end{aligned}$$

x dipende dall'ingresso u (e l'ingresso u deve essere moltiplicato per i coefficienti relativi alla parte MA, b), più le uscite ai passi precedenti che vengono moltiplicate per i coefficienti AR (a). Questa relazione è nota dalla teoria dei filtri digitali, solo che nella teoria dei filtri digitali l'avevamo vista scritta così:

Slide 8-9-10: PSD DI UN MODELLO LTI (ARMA) – PSD DI MODELLI MA E AR

Siccome il mio obiettivo è quello di calcolare la PSD dell'uscita (P_{xx}), a questo punto si può scrivere la PSD in funzione della PSD dell'ingresso, moltiplicato il modulo quadro della funzione di trasferimento, cioè si sta utilizzando la proprietà dei sistemi LTI per cui lo spettro dell'uscita è uguale allo spettro dell'ingresso per il modulo quadro della funzione di trasferimento:

$$P_{xx}(f) = P_{uu}(f) \cdot |H(f)|^2$$

Questo vale nel dominio di Fourier, mentre nel dominio zeta cambia solo che al posto di $|H(f)|^2$, c'è $H(z)$ per il suo complesso coniugato:

$$P_{xx}(z) = P_{uu}(z) \cdot H(z) \cdot H^*(1/z^*)$$

Siccome:

$$H(z) = \frac{B(z)}{A(z)}$$

Allora si può scrivere anche in questo modo:

$$P_{xx}(z) = P_{uu}(z) \cdot \frac{B(z)B^*(1/z^*)}{A(z)A^*(1/z^*)}$$

Per cui si può calcolare la PSD del nostro segnale, direttamente da una delle due formule di $P_{xx}(z)$ => questo vuol dire che è stato dimostrato che la forma dello spettro in uscita, è dato da quello che ci mette all'ingresso per la funzione di trasferimento. Allora si capisce che usare come sequenza di ingresso $u[n]$, un processo casuale gaussiano bianco, ci facilita le cose perché lo spettro di un processo causale gaussiano bianco è una costante (e questa costante è ρ_u , la varianza di quel processo casuale).

Quindi alla fine si scopre che la PSD del nostro segnale è:

$$P_{ARMA}(f) = T\rho_u |B(f)/A(f)|^2$$

ovvero il prodotto di una costante per la funzione di trasferimento del modello. Abbiamo risolto il punto 4), cioè è stato dimostrato che una volta che abbiamo il modello, possiamo calcolare la PSD a partire dal modello; ma un modello dipende dai parametri, e quindi quella PSD la chiamiamo PSD PARAMETRICA.

In realtà, in questo caso, si sta considerando una funzione di trasferimento $H(z)$ che ha sia la parte MA che quella AR e quindi ho ricavato una PSD PARAMETRICA derivante da un modello ARMA, perché all'interno del modulo quadro ci sono sia poli che zeri.

Abbiamo detto che se utilizzassimo un modello MA, avremmo solo il polinomio al numeratore, cioè avremmo solo i coefficienti b e allora una PSD parametrica derivante da un modello MA si scrive così:

$$P_{MA}(f) = T\rho_u |B(f)|^2$$

mentre una PSD parametrica derivante da un modello AR si scrive così:

$$P_{AR}(f) = \frac{T\rho_u}{|A(f)|^2}$$

Quindi, una volta che abbiamo il modello, calcolare la PSD è banale: si prende il modulo quadro della funzione di trasferimento. Il problema è che non abbiamo il modello => il difficile è crearsi quella $H(z)$. Però abbiamo detto che ci sono 3 modelli, bisogna scegliere quello giusto e poi bisogna calcolare a e b (o solo a o

tradizionale, ad ogni frequenza si legge esattamente il valore di potenza associato => bisogna iniziare a capire che i 2 spettri presentano informazioni leggermente differenti).

• **PASSO 1: SCELTA DEL MODELLO**

Slide 12: RELAZIONE TRA MODELLI ARMA, MA E AR

A questo punto bisogna scegliere il modello tra: ARMA, AR e MA. Possiamo notare che scegliere un modello AR piuttosto che MA o piuttosto che ARMA, alla fine, non è una cosa abbastanza sconvolgente; infatti il primo passo (quasi mai) causa dei grossi problemi, sia perché le scelte sono limitate e sia perché alla fine si può dimostrare matematicamente che qualunque funzione di trasferimento $H(z)$ si può ottenere sia da un modello AR, che da un modello MA, che da un modello ARMA. Quello che cambia macroscopicamente è l'ordine => anche in questo caso si può dimostrare che la stessa $H(z)$ si può ottenere quasi identica da qualunque modello, di ordine diverso. Allora purchè noi adattiamo l'ordine del nostro modello, la scelta di uno o dell'altro, non è particolarmente significativa. La scelta ricadrà inevitabilmente sempre e solo su modelli AR => il punto 1), ovvero la scelta di un modello, non è così problematico per questi motivi:

- si può scegliere quello che ci pare, tanto alla fine (se lo si fa bene), si riesce ad ottenere in ogni caso una descrizione buona dello spettro del tuo segnale;
- dal punto di vista pratico, alla fine, si userà sempre un modello AR.

Perché si userà sempre e solo un modello AR? Il punto di partenza è questa equazione:

$$\begin{aligned}
 x[n] &= -\sum_{k=1}^p a[k]x[n-k] + \sum_{k=0}^q b[k]u[n-k] \\
 &= \sum_{k=0}^{\infty} h[k]u[n-k]
 \end{aligned}$$

non abbiamo né a, né b, né u ma abbiamo solo x e la descrizione di x tramite il modello => questo è il punto di partenza, ed è l'equazione del modello ARMA. Noi però sceglieremo sempre un modello AR perché il problema è che l'unica cosa che noi abbiamo è x, tutto il resto non lo abbiamo. Questo sistema è un sistema lineare di ordine k, perché è la descrizione di un polinomio e abbiamo visto che l'equazione temporale di un filtro si può scrivere in forma ricorsiva: scrivendola in forma ricorsiva, se il filtro è di ordine n, abbiamo n equazioni tutte collegate una all'altra => questo è un sistema lineare di ordine k, dove se io voglio descrivere lo spettro, ho bisogno di calcolare $H(z)$, ma a sua volta $H(z)$ è determinata dai coefficienti a per un modello AR (oppure dai coefficienti a e b se fosse un modello ARMA). Allora se devo risolvere questo sistema lineare, contando che ho x e quindi l'incognita è a, cioè sono i parametri => siccome u non lo conosco, non potrò mai calcolare b; al più ci posso mettere un numero a caso $u[n]$ (uno solo) e poi risolvere il sistema per calcolare i coefficienti a, che dipendono da x (ma x lo conosco).

Quindi per questioni di lavorabilità algoritmica, l'unico modello che riusciamo a gestire, è il modello AR perché lega i parametri al segnale stesso, che è l'uscita di quel modello => preferiamo lavorare con modelli AR perché c'è il beneficio algoritmico che quel sistema, che poi ci porterà al calcolo dei parametri, è un sistema che per noi è gestibile altrimenti diventerebbe completamente ingestibile.

Slide 13-14: MODELLO AR DI ORDINE INFINITO E CALCOLO PARAMETRI

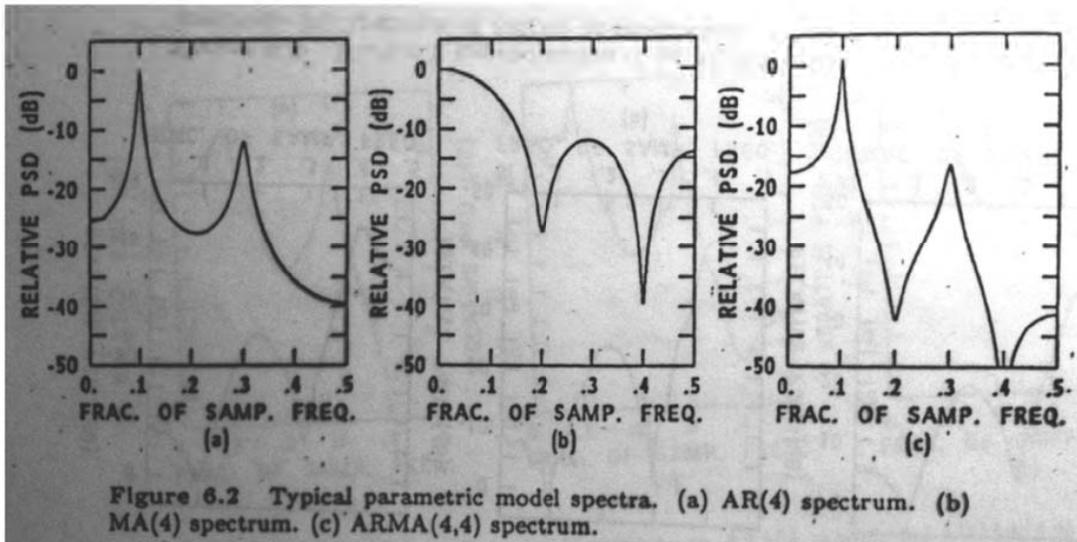
Se scelgo un modello di tipo diverso non ci sono problemi, nel senso che si può dimostrare che la stessa funzione di trasferimento la possiamo descrivere come ci pare, con modelli di qualunque tipo. Per dare una evidenza matematica, si parte da una cosa di questo tipo: supponiamo di avere un modello AR di ordine infinito e supponiamo che i coefficienti di questo modello AR si chiamino c. Allora scriviamo un generico modello AR di ordine molto grande, che tende ad infinito, che ha un vettore di coefficienti c => il polinomio

a cui da origine, è la trasformata zeta di questi coefficienti c: $C(z) = 1 + \sum_{k=1}^{\infty} c[k] z^{-k}$.

Quindi noi abbiamo visto il caso generico di avere un numero infinito di coefficienti c ; in realtà se c è in numero finito, il sistema lineare ha un ordine finito e quindi possiamo risolverlo per qualsiasi ordine \leq .

Slide 15: FUNZIONI DI TRASFERIMENTO DEI VARI MODELLI

Nella pratica questa non è una limitazione più grave di tanto, anche perché gli spettri hanno degli aspetti un po' diversi. Per esempio supponiamo di avere un modello AR di ordine 4, con le radici messe in qualche punto => con questo modello, se andiamo a fare il modulo quadro della funzione di trasferimento, ottengo una roba fatta come in figura (a):



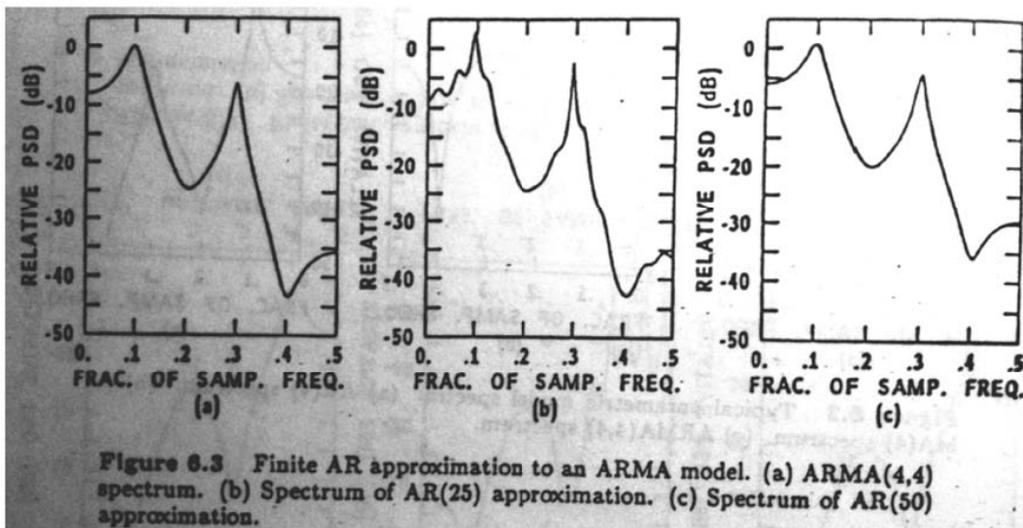
I modelli AR generano sempre delle funzioni di trasferimento caratterizzate da picchi stretti e valli ampie.

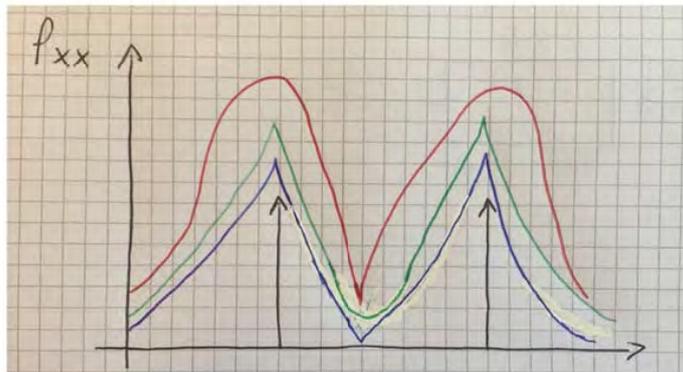
Se invece prendo il modello MA sempre di ordine 4, lo spettro è quello in figura (b) => è il contrario di quello AR, perché si hanno picchi ampi e valli strette.

Un modello ARMA di ordine (4,4), ha invece una PSD fatta come in figura (c): picchi stretti e valli strette.

Slide 16: APPROSSIMAZIONE AR DI UN FILTRO ARMA

Bisogna fare un'altra considerazione: la slide precedente ci dice una cosa abbastanza facile, ovvero che ogni modello genera una PSD con delle caratteristiche specifiche.





Allora bisogna fare attenzione, perché se dovessimo andare a calcolare la potenza associata a quell'armonica o associata ad una determinata banda, prima si faceva l'integrale, ma fare l'integrale di una roba fatta così (BLU), piuttosto che di una funzione fatta così (ROSSO), è molto diverso, perché l'area è diversa. Quindi a parte per il fattore di scala che c'è davanti (ρ_u che non conosciamo), il modello influenza anche l'area che noi andiamo a calcolare quando facciamo l'integrale => è per questo che in precedenza abbiamo detto che dagli spettri parametri calcolare la potenza, è un po' complicato perché varia a seconda dell'ingresso e a seconda del modello scelto; quindi funziona meglio quando dobbiamo calcolare delle potenze relative, ma funziona meno bene quando dobbiamo calcolare delle potenze assolute.

NB: per i nostri segnali, non è un problema troppo limitante, però ci possono essere dei casi dove bisogna fare attenzione.

Slide 18÷22: DAI PARAMETRI DEL MODELLO ALLA r_{xx} [m]

Abbiamo visto che esiste una relazione tra i coefficienti e i modelli di tipo diverso e di ordine diverso, per cui si può saltare "agevolmente" da un modello all'altro, calcolando di volta in volta i parametri del modello che ci interessa, purché mettiamo d'accordo l'ordine.

Adesso vediamo che i parametri del modello, cioè i coefficienti, sono legati non solo tra di loro, ma anche agli elementi della funzione di autocorrelazione del segnale (che ci servirà in seguito per introdurre una caratteristica della stima spettrale).

Partiamo dalla solita relazione:

$$x[n] = -\sum_{k=1}^p a[k]x[n-k] + \sum_{k=0}^q b[k]u[n-k]$$

che è la descrizione del modello LTI di ordine p,q.

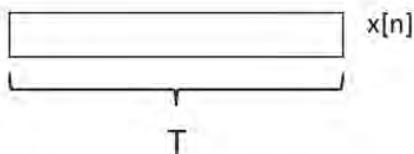
Adesso moltiplichiamo entrambi i membri, per: $x^*[n-m]$

E poi di tutto ne calcoliamo il valore atteso, cioè:

$$E\{x[n]x^*[n-m]\} = -\sum_{k=1}^p a[k]E\{x[n-k]x^*[n-m]\} + \sum_{k=0}^q b[k]E\{u[n-k]x^*[n-m]\}$$

Questa roba qua: $E\{x[n]x^*[n-m]\}$ è la sequenza di autocorrelazione (ACS) r_{xx} [m], cioè il valore atteso di quel prodotto, da' origine ad un insieme di elementi che sono gli elementi della ACS. Si chiama r_{xx} [m] perché m è il ritardo che c'è tra il segnale e la sua replica.

risoluzione spettrale che non potevamo ottenere nella pratica. Allora la vera stima parametrica operativa, parte da qua, con una considerazione molto semplice: eravamo partiti da una situazione di questo tipo

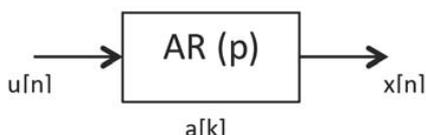


$x[n]$ è il nostro segnale, con un supporto temporale limitato pari a T . Con la stima spettrale tradizionale, la risoluzione spettrale teorica che mi posso permettere è $\Delta f = \frac{1}{T}$. Per qualche motivo, questa risoluzione spettrale teorica è inadeguata allo spettro che devo descrivere (era il caso dei segnali di ossigenazione che abbiamo visto..) => non posso aumentare T , perché il segnale o diventa non stazionario, oppure non c'è, cioè c'è un segnale biologico a un potenziale, che dura T secondi e quindi il resto non esiste. Se da questa $x[n]$ calcolo la sua funzione di autocorrelazione:

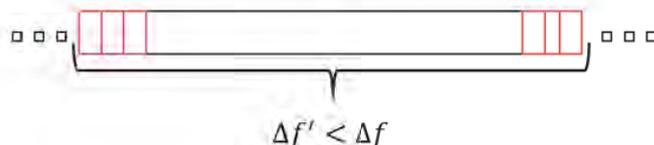


anche questa funzione di autocorrelazione ha una lunghezza limitata (si è visto che se ho un certo numero di campioni, e continuo a calcolare i ritardi oltre qual numero massimo di elementi del segnale, tutti i coefficienti della ACS sono nulli) => la lunghezza T si ripercuote sulla lunghezza della ACS. Ovviamente questo è proprio quello che fa sì che Δf , sia finita.

Se ho adesso un modello LTI, che è un modello AR di ordine p , in cui ci entra quello che mi pare ed esce il mio segnale $x[n]$, e questo modello è un bel modello, ho un certo numero di coefficienti autoregressivi: ne ho $p+1$.



Questa equazione $r_{xx}[n] = -\sum_{k=1}^p a[k]r_{xx}[m-k]$, ci dice che da questi coefficienti AR, gli do in pasto la ACS e lui consente di calcolare elementi della ACS di ordine successivo. Vuol dire che gli do in pasto la ACS e lui mi permette di calcolare altri elementi della ACS, che prima non avevo => la ACS diventa più lunga. Ma come abbiamo visto, la lunghezza della ACS è direttamente correlata alla risoluzione spettrale e questo vuol dire che se adesso calcolo la risoluzione spettrale $\Delta f'$, ottenibile da questa ACS allungata, ovviamente si ha che: $\Delta f' < \Delta f$, cioè vuol dire che la risoluzione spettrale teorica è aumentata.



Quindi ancora non sappiamo come si calcola quella funzione autoregressiva di ordine p e quali sono i coefficienti a che mi consentono di fare ciò, però supponendo di avere i coefficienti a , perché funziona la stima spettrale parametrica, c'è scritto qui, perché dopo che ho modellizzato il segnale, su quel modello ho una risoluzione spettrale teorica di quanto mi pare. Quanti ne posso calcolare di quei coefficienti?? Infiniti,

04/05/2015

Riepilogo...

Stima spettrale parametrica vuol dire approccio modellistico, il che si significa sistema LTI in cui il nostro segnale è l'uscita e l'ingresso è un rumore casuale, che per comodità consideriamo gaussiano bianco. La funzione di trasferimento di questo sistema LTI è un modello AR (perché abbiamo visto che è più comodo da trattare), ma potrebbe essere benissimo un modello qualunque tanto parametri di modelli diversi, sono collegati tra di loro dalle equazioni viste. Inoltre abbiamo scoperto che i parametri di un generico modello AR hanno anche una relazione con gli elementi della ACS del segnale stesso. Quindi il problema adesso è: abbiamo scelto il tipo di modello, dobbiamo selezionare l'ordine.

Slide 24-25: PREDITTORE LINEARE FORWARD E PREDITTORE LINEARE COME MODELLO AR

In particolare abbiamo visto il predittore lineare, che non è altro che una combinazione lineare di elementi per originare elementi successivi, dove i pesi della combinazione lineare, sono gli elementi del modello AR considerato. La capacità di legare i coefficienti del modello AR agli elementi della ACS, la possiamo usare non solo per allungare la ACS, ma anche applicata direttamente al segnale (come effettivamente si usa), cioè definiamo un predittore lineare che ha questa struttura:

$$\hat{x}^f[n] = -\sum_{k=1}^m a^f[k]x[n-k]$$

Il cappellino indica che è un elemento stimato (non reale) => l'elemento della sommatoria di posto n è uguale a meno la sommatoria di k che va da 1 a m dei coefficienti a[k] per x[n-k]. Questa struttura è uguale a quella alla slide 22, solo che al posto di r_{xx} , c'è x e al posto di n c'è m (ma la struttura è la stessa).

=> Si chiama predittore lineare perché sulla base di un certo numero di campioni precedenti, stimiamo il campione futuro e lo chiamiamo *predittore lineare diretto* (la f ad apice indica *forward*, cioè *diretto*) perché vuol dire che dati i campioni n-qualcosa, calcolo il campione di posto n-esimo. In realtà teoricamente è stato dimostrato che ci potrebbe essere anche un predittore inverso, cioè dati i campioni futuri, calcola i campioni precedenti (potrebbe non avere senso perché prima ho i campioni precedenti e poi quelli futuri, ma la direzione dell'asse del tempo è una convenzione mia).

Questa descrizione è simile alla descrizione di un modello AR, perché servono gli stessi campioni precedenti di un segnale per generare il campione al passo successivo. Non è proprio un modello AR, perché se fosse tipo un modello AR come quello visto nei filtri, ci vorrebbe un ingresso da qualche parte, mentre qui non c'è; qui è proprio x che si ripete e genera se stesso, non c'è un ingresso u o qualcosa del genere => ha la struttura di un modello AR, ma non è un filtro AR o la funzione di trasferimento AR, come normalmente conosciamo. La cosa che dobbiamo ricordare è che il numero di questi coefficienti è finito, in particolare ce ne sono m, che è "l'ordine" del filtro, che è esattamente quello che noi adesso andiamo a cercare, cioè si è capito che c'è un collegamento tra i parametri di questo modello e il mio segnale, e quindi possiamo sfruttare questo collegamento per andare a ricavare un po' di informazioni utili.

La prima è stabilire l'ordine: una volta stabilito il modello, dobbiamo stabilire l'ordine. Allora quello che in questo momento noi cerchiamo è tentare di capire quale è l'ordine di questo modello che ci va bene. Non è una cosa da poco, perché la sommatoria, dato che va da 1 a m, fa sì che io abbia bisogno di m valori precedenti per calcolare il successivo => vuol dire che, per esempio, se m=20 e ho un segnale lungo 100milioni di campioni, non mi servono tutti i 100milioni, me servono 20 per calcolare quello dopo. Allora si capisce che fare una predizione su 20 campioni piuttosto che su 30,40 o 100, è una cosa diversa. Allora la domanda che ci poniamo è: ammettiamo che funzioni questo procedimento, di quanti campioni ho bisogno per fare una predizione ragionevole del mio segnale?? Quindi il problema principale è capire quanti campioni devo scegliere per predire in modo ragionevole quello che esce.

abbiamo visto che dal punto di vista numerico, è l'unico che possiamo trattare per le equazioni, perché l'uscita dipende da se stessa (e noi abbiamo solo x , non abbiamo niente di tutto il resto). Supponiamo che l'ordine di questo modello AR, sia pari a p e quindi ovviamente c'è una serie di coefficienti a di questo modello per k che va da 0 a p (quindi ce ne sono $p+1$, dove il primo vale sempre 1, ecc...). In uscita il nostro segnale che stiamo considerando, ha una certa PSD => stiamo partendo da un ingresso che ha una PSD piatta, troviamo in uscita una PSD che è quello che ci pare e ovviamente a forma gliel'ha data la funzione di trasferimento di quel sistema LTI. Oggi abbiamo detto che mettiamo un altro sistema LTI che è il predittore lineare => dalla struttura che abbiamo visto, il predittore lineare ha una struttura che è simile a un modello AR di ordine m . Da quella struttura, succede che questo modello prende in ingresso la mia sequenza numerica e produce in uscita quello che è il mio errore di predizione. Qui non facciamo tanta matematica, ma solo un sforzo di intuizione: supponiamo di andare a calcolare la PSD del nostro errore di uscita e supponiamo di trovare che la PSD dell'errore di predizione, in uscita dal nostro predittore lineare, è piatta. Detto in altre parole, l'errore di predizione è un rumore gaussiano bianco.

Allora abbiamo un po' chiuso il cerchio, perché abbiamo detto: partiamo da un rumore gaussiano bianco, lo filtriamo con una funzione di trasferimento che gli dà una forma, poi lo filtriamo al contrario (perché nel modello sopra $x[n]$ è l'uscita mentre nel modello sotto è l'ingresso) con una funzione di trasferimento che la forma gliela toglie e dà in uscita uno spettro piatto: piatto era all'inizio ($u[n]$) e piatto è alla fine => vuol dire che le due funzioni di trasferimento, fanno a stessa cosa. Questo vuol dire che se una mi fa passare da bianco ad un certo spettro e l'altra mi fa passare da un certo spettro a bianco, sono esattamente la stessa cosa, solo invertite una rispetto all'altra. Allora vuol dire che se ciò avviene, questi due modelli AR di ordine p e AR di ordine m , coincidono. Quindi succede che, a differenza del modello AR tradizionale, la sequenza di errore è l'uscita del predittore lineare diretto (con $x[n]$ che è l'ingresso) e succede che se l'errore è rumore gaussiano bianco, ovviamente vuol dire che l'errore normalmente è scorrelato dall'ingresso, cioè da quello che noi ci mettiamo nel predittore lineare. In particolare quando è un rumore gaussiano bianco, abbiamo che questo modello AR (che abbiamo descritto nel nostro predittore lineare) coincide con quello che ha generato il segnale.

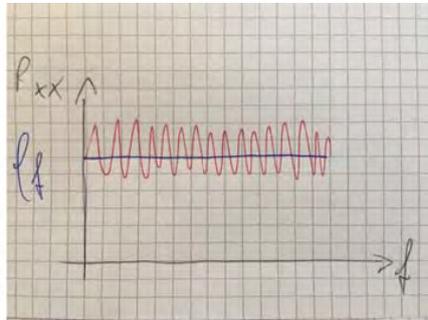
In pratica: conosco un segnale ma non riesco a farne stima spettrale per i motivi detti, allora lo modellizzo e se suppongo che il segnale è stato costruito da un processo AR, quando mi trovo, al contrario, un processo che mi dà in uscita un errore di predizione che non contiene informazione spettrale (perché è una costante) => vuol dire che quel modello si è preso tutta l'informazione del mio segnale e quindi vuol dire che lo spettro di quel modello (cioè la funzione di trasferimento di quel modello) coincide con la funzione di trasferimento di chi l'ha creato. In particolare vuol dire che l'ordine m è uguale all'ordine p => siccome i passi sono stati:

- scegliere un modello
- determinare l'ordine
- e poi calcolare i parametri

=> il modello l'ho scelto (autoregressivo); l'ordine lo calcolo prendendo la sequenza numerica, dandola in pasto al predittore lineare e quando l'errore di predizione è una sequenza bianca, vuol dire che sono arrivato all'ordine giusto => fintanto che l'ordine è più basso, l'errore di predizione non sarà una sequenza di rumore gaussiano bianco, ma sarà una sequenza di rumore, scorrelato dal predittore, ma non bianco. Quindi arrivo ad avere in uscita uno spettro piatto solo quando $m=p$. Questo predittore che prende in ingresso una sequenza casuale qualunque (il nostro segnale $x[n]$) e ne dà in uscita un errore casuale (un segnale casuale bianco) si chiama *filtro sbiancante* -> si chiama filtro sbiancante perché quando ho uno spettro che ha una certa sagoma, in uscita mi ritrovo un errore che ha uno spettro piatto, vuol dire che tutta l'informazione è rimasta intrappolata all'interno del filtro. Quello che si ha in uscita dal punto di vista spettrale (un rumore gaussiano bianco) ha un'informazione nulla, l'unica cosa

fattore di qualità, Q) → più lo stimatore diventa consistente, più quelle oscillazioni crescono e più difficilmente si può dire che quello sia bianco o no.

2) Guardare come è fatto lo spettro → dal punto di vista teorico mi aspetterei una roba così:



dove ρ^f è la varianza di rumore. In realtà non ottengo una roba così (BLU), ma ottengo una roba (ROSSO), come già visto, che è molto variabile perché dipende da che stimatore uso, da come lo uso, ecc.. in realtà il problema è la consistenza o il fattore di qualità. È difficile prendere una decisione, perché dovrei sapere se lo spettro è costante oppure no: certo che non è costante, ha delle oscillazioni casuali, ma il problema sarebbe quello di capire se le oscillazioni sono contenute in una fascia ragionevolmente stretta per cui si può dire che sono effettivamente oscillazioni casuali e non è contenuto del segnale. Anche questo dipende dal tipo di stimatore, dal numero di campioni, dalla finestra e dalla consistenza, cioè dal fattore di qualità.

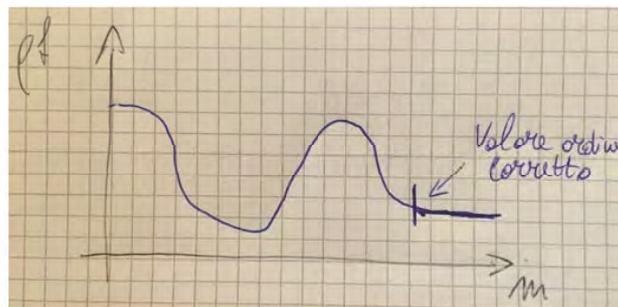
Quindi in realtà i metodi per sapere se una sequenza numerica è assimilabile a un rumore bianco, sono 2: o si va a vedere come è fatto lo spettro o si va a vedere come è fatta la ACS → è la stessa cosa, perché tra una e l'altra c'è di mezzo la trasformata di Fourier. Entrambi i metodi hanno delle limitazioni e in pratica sono dei metodi che dal punto di vista matematico sono inconsistenti. Come fanno i matematici a sapere se una serie numerica è effettivamente una serie numerica gaussiana bianca oppure no? I matematici usano altri metodi, molto più complicati (che per il momento non vediamo), ma che sono molto efficienti nel senso che ci sono degli operatori particolari che, se applicati ad una sequenza bianca, producono in uscita zero => ci sono degli indicatori particolari che per i rumori bianchi sono effettivamente nulli (e quindi si fa alla svelta a sapere se una sequenza è bianca oppure no). Per noi invece dipende da come si fanno le cose e da che decisione si prende: è inevitabile che se prendiamo la stessa sequenza numerica (bianca o non bianca che sia) e la processa con determinati parametri e prendo una decisione e poi la processa qualcun altro, magari si hanno 2 esiti diversi, perché dipende da troppa cose per avere un'indicazione consistente. Quindi nella pratica, avere queste limitazioni, vuol dire che questi metodi noi difficilmente li possiamo usare fidandoci, però per noi è una cosa cruciale, perché stabilire quando una sequenza numerica è bianca, vuol dire aver trovato l'ordine giusto del nostro modello (ricollegandoci alla stima spettrale). Infatti sapere dire se una sequenza numerica è bianca oppure no, vuol dire aver trovato l'ordine corretto che per noi è uno dei passaggi fondamentali.

Slide 29-32: COINCIDENZA DEI DUE MODELLI E VARIANZA DELL'ERRORE DI PREDIZIONE

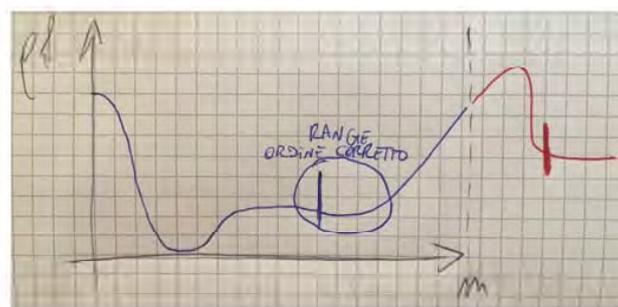
Questo metodo è il metodo teoricamente corretto → per cui se ci chiedono: "qual è il metodo teorico per sapere quando il modello ha l'ordine giusto??" la risposta è "quando l'errore di predizione è bianco". Ma nella pratica non saremo mai capaci di farlo (è molto difficile farlo) => non si applica tendenzialmente un metodo teorico, ma si applica un metodo pratico, empirico, che si basa una considerazione successiva. Si è detto che quando l'ordine è corretto, il nostro ordine m coincide con l'ordine p del modello AR che ha generato il segnale. Ovviamente i coefficienti coincidono e le varianze coincidono. Quindi:

Questo siamo capaci di farlo e vedremo che calcolare la varianza dell'errore, a partire dall'ordine, è talmente facile che non c'è nemmeno bisogno di risolvere tutto il predittore lineare, è una roba molto veloce. Allora se ci viene dato un segnale x , quello che sicuramente si può fare è disegnare il diagramma della varianza per un numero di ordini variabile (variabile tra quanto e quanto, si vedrà).

NB: cosa che succederà durante l'esercitazione → la teoria dice che la varianza dell'errore, ρ^f , è monotona decrescente al crescere dell'ordine del modello (o del filtro, è la stessa cosa); nella pratica però, monotona decrescente non lo è mai, nel senso che è facile che abbia delle oscillazioni per cui non sia monotona decrescente. Però bisogna distinguere un certo numero di casi → ad esempio supponiamo di avere un diagramma della varianza dell'errore fatto così:



l'ordine corretto non è il minimo della varianza dell'errore, ma è quando la varianza dell'errore diventa asintotica. Teoricamente questo diagramma non è corretto perché un predittore lineare non dovrebbe dare una varianza dell'errore fatta così, perché dovrebbe essere monotona decrescente (però nella pratica succede per tutta una serie di questioni, a partire da quella più ovvia che sono le approssimazioni numeriche). Ma ce n'è un'altra → supponiamo che la varianza dell'errore sia fatta in quest'altro modo:



l'ordine corretto qual è? Qui ci sono 2 questioni da tenere in considerazione: potrebbe anche essere che la varianza dell'errore, a un certo punto fa così (ROSSO) e quindi in realtà l'ordine giusto è questo (LINEA ROSSA) → cioè potrebbe anche dire che noi abbiamo calcolato la varianza dell'errore per un numero di ordini elevato (difficile però per i nostri segnali, perché gli ordini non sono mai elevatissimi, al massimo abbiamo fatto filtri di ordine 50-51). Ragionevolmente, allo stesso modo, l'ordine massimo del predittore lineare non sarà 1500, non sarà 200, ma sarà qualche decina => avremo predittori lineari con ordine, per i nostri segnali biologici, che sta tra una decina e qualche decina. Allora se supponiamo che il range che ci interessa, fino a qualche decina, è qua (LINEA TRATTEGGIATA NERA), dobbiamo ricordarci che stiamo andando a cercare l'ordine di quel modello AR uguale all'ordine del modello AR che ha generato il segnale. E se il segnale non è stato generato da un modello AR? Cioè, noi stiamo facendo un'hp di base che il nostro segnale $x[n]$ deriva da un processo AR, ma nella pratica il nostro segnale biologico non deriverà mai da un modello AR (alcuni segnali sì, perché sono

giusto; stabilito l'ordine giusto, quel valore di m , lo do in pasto all'algoritmo e calcolo i parametri solo per quell'ordine lì => così calcolo i parametri del modello, ma ancora devo fare stima spettrale.

Questo metodo, che è implementato, serve sia per calcolare il diagramma di varianza e sia per calcolare i parametri; in realtà però non è l'unico → quando si considera un modello AR di ordine m si stabilisce solo che ci sono $m+1$ coefficienti e questi $m+1$ coefficienti si possono calcolare in mille modi diversi. L'algoritmo di Levinson li calcola a partire dagli elementi della ACS del segnale.

Slide 35-38: CRITERI POSSIBILI PER IL CALCOLO DEI PARAMETRI

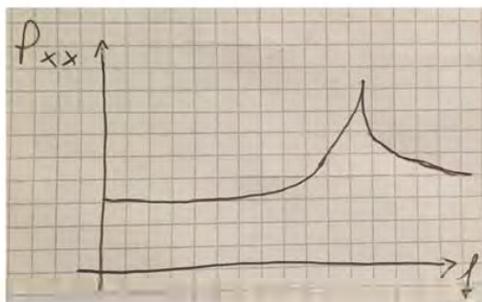
Se prendo questo set di istruzioni e le risolvo, sto calcolando le radici secondo quello che viene chiamato metodo di Yule-Walker, per cui i coefficienti vengono calcolati direttamente risolvendo l'algoritmo di Levinson e però trovo un problema che non mi piace molto: normalmente questo metodo di calcolare i parametri non si usa molto perché mi calcola dei parametri che mi restituiscono uno spettro troppo "liscio", non ho abbastanza risoluzione (soprattutto quando x è una sequenza numerica molto corta, cioè fa molta fatica su segnali molto corti).

Ci sono 3 metodi alternativi a questo (tutti implementati in Matlab):

- 1) Il *metodo di Burg* (detto anche criterio a massima entropia) risolve lo stesso sistema lineare di prima (leggermente modificato) ma invece di calcolare passo-passo tutti i coefficienti, calcola tutti i coefficienti in una volta sola, massimizzando l'entropia dei coefficienti.

L'entropia, che è una misura di probabilità, ci dice che dato un sistema, la probabilità di avere la stessa uscita è molto bassa => è un sistema "disordinato", ma in realtà concettualmente è una misura di probabilità. Allora massimizzare l'entropia vuol dire trovare una soluzione in modo che quel sistema lineare dia in uscita valori che hanno una probabilità molto bassa di ripetersi uguali (anche a distanza di tempo molto grande). Se ci viene data una sequenza numerica e dobbiamo calcolare l'entropia di quella sequenza numerica, non c'è problema perché sapendo che è una misura collegata in qualche modo alla probabilità, la prima cosa da fare è calcolare l'istogramma (o la funziona PSD: sono la stessa cosa perché l'istogramma è la versione discreta della PSD) e calcolare ogni valore numerico quante volte si ripete → se la probabilità che si ripeta tante volte è alta (cioè si hanno colonne molto alte), allora l'entropia sarà bassa; se invece si ha un istogramma basso, molto lungo, con dei valori molto piccoli, allora l'entropia sarà alta.

La stessa cosa succede qui: l'ordine è fissato, ma per posizionare le radici bisogna fare in modo che nel rumore di uscita ci sia un'entropia massima possibile. Però bisogna fare attenzione con questo metodo, perché specie per ordini molto alti, si può introdurre bias nella stima di frequenze sinusoidali. Cosa vuol dire? Supponiamo che in frequenza lo spettro reale sia questo (modellizzato con il nostro processo AR):



lo spettro è corretto, ma la frequenza è spostata da una parte. Allora quello che può capitare con il metodo di Burg (specie per ordini molto alti), le linee spettrali si spostano, non sono centrate rispetto a quelle originali.

Slide 39-40: STIMA SPETTRALE AD ALTA RISOLUZIONE ED ESEMPIO

Calcolati i valori dei parametri, si calcola lo spettro → sulla base dei parametri, basta calcolare: $\frac{1}{\text{polinomio che deriva dai parametri AR}}$ e si ha la misura dello spettro del segnale, che per il teorema di W-K, ovviamente, è uguale allo spettro del segnale reale.

Questa relazione:

$$P_{AR}(f) = \frac{T \rho_u}{|1 + \sum_{k=1}^p a[k] e^{-j2\pi k f T}|^2} = T \sum_{k=-\infty}^{+\infty} r_{xx}[k] e^{-j2\pi k f T}$$

fa vedere come c'è un legame tra i coefficienti del modello AR e gli elementi della ACS.

Come detto, lo spettro del segnale è sempre quello, sia che lo calcoli con il teorema di W-K (ovvero come trasformata di Fourier della ACS) o spettro AR tramite il predittore lineare → bisogna ottenere la stessa cosa. Ovviamente però l'aspetto è diverso perché si hanno vincoli di risoluzione diversi: W-K impone un vincolo di risoluzione che lega la lunghezza della ACS alla risoluzione spettrale, mentre questo modello no e quindi si può calcolare la funzione di trasferimento di questo modello (che è una quantità continua) come voglio io, anche con pochi coefficienti. Quindi l'unica differenza è questa qua, però lo spettro, quello reale, è diverso (concettualmente). Allora qualcuno ha detto: se uso gli elementi della ACS per calcolare i parametri del predittore lineare (algoritmo di Levinson), a un certo punto gli elementi della ACS non se ne hanno più e quindi al massimo calcolo un certo numero di coefficienti AR. Una volta che ho questo modello e che questi due spettri coincidono, si è visto però che si può allungare la ACS fittiziamente, in modo da aggiungere altri elementi (che sono elementi predetti) e che quindi fanno sì che la risoluzione temporale, dalla parte destra (cioè teorema di W-K), aumenta perché aumenta la lunghezza della ACS. E quindi è stata introdotta quella che si chiama stima spettrale ad alta risoluzione: sugli elementi veri della ACS, calcolo i parametri del modello; poi dai parametri del modello, allungo la ACS (con la relazione vista la scorsa volta) => mi ritrovo con una ACS più lunga che prima non avevo e quindi con una risoluzione spettrale maggiore. Teoricamente posso arrivare ad una risoluzione spettrale infinita, perché di parametri della ACS ne posso calcolare quanti ne voglio. Quindi è chiamata stima spettrale ad alta risoluzione o stima spettrale a risoluzione infinita → in entrambi i casi è una cazzata però perché è una risoluzione apparente, non è una risoluzione reale: tutte le volte che si aggiungono elementi interpolati, non si sta aggiungendo informazione vera, ma informazione di comodo (non è vietato fare ciò, anzi ci sono casi in cui è un bene farlo).

se il segnale non è più stazionario, non posso più scrivere $r_{xx}[m]$, ma devo scrivere $r_{xx}[m,n]$, perché varia nel tempo. Come si fa a modellizzare uno spettro che nel tempo cambia?! Non ci riesco, non si può modellizzare qualcosa che varia nel tempo. Inoltre, dato che in ingresso ho un segnale u che è certamente stazionario, non si può avere in uscita un segnale x non stazionario. Al passo n -esimo, al primo campione, si ha una certa funzione di autocorrelazione e calcolo un parametro a ; al passo successivo, cambia la funzione di autocorrelazione, perché il segnale non è stazionario e quindi la funzione di autocorrelazione varia nel tempo => il secondo parametro a che calcolo, è scorrelato dal primo. Quindi tutte le volte che bisogna modellizzare qualcosa, ad esempio lo spettro, questo deve stare fermo → se varia non c'è modo.

Esistono anche approcci che vanno bene anche per segnali che variano nel tempo, ma non si chiamano AR, ma si chiamano ARV o ARVx, ad indicare il fatto che sono in grado di modellizzare anche spettri che hanno delle variazioni temporali.

Questo dipende dal fatto che l'ordine determina il numero di radici → il modello AR, cioè il predittore lineare, le radici le posiziona sull'asse delle frequenze in modo da ottimizzare l'informazione, cioè da fare in modo che quello che esce dal predittore lineare sia un rumore bianco (cioè tenta di sbiancare il segnale). Una volta che l'ha sbiancato, in teoria tutte le radici successive sono 0, nella pratica no perché il segnale difficilmente sarà AR; quindi se si alza l'ordine del modello, non è che si avranno tutte le radici nulle, ma qualche radice sarà nulla e qualche radice non sarà nulla. Siccome però l'uscita è già bianca, a questo punto le radici si posizionano a caso, e quindi a seconda di dove si posizionano, vanno a finire sull'asse delle frequenze e possono dare origine a "sparate" che non hanno niente a che vedere con lo spettro del segnale.

2) Altra cosa che può succedere è il *bias di componenti sinusoidali* → più facile quando calcoliamo i coefficienti con il metodo di Burg, meno facile con gli altri metodi. Quindi vuol dire che una riga spettrale invece di essere dove dovrebbe essere, è spostata.

3) La terza cosa è lo *splitting di righe spettrali* → vuol dire che se questo (NERO) è come dovrebbe essere fatto teoricamente uno spettro, alzando l'ordine succede una roba di questo tipo (BLU), ovvero la riga si divide in due:



Quindi aggiungendo tante radici, le mette dove riesce e quindi a questo punto è facile che, invece di avere un bel picco, si divide in due → questo è quello che si intende per splitting di righe spettrali.

Altri metodi per capire se l'ordine è corretto oppure no, osservando lo spettro, non ce ne sono, tranne una considerazione ovvia: supponiamo di avere un segnale, di modellizzarlo con un modello AR e supponiamo che l'ordine corretto di questo modello sia ad esempio 20. Se osservo lo spettro per un ordine molto più basso di 20, tipo 5, 6 o 7, ovviamente sarò nella prima condizione, cioè vedo uno spettro molto blando che non mi soddisfa. Man mano che ci avviciniamo all'ordine, succede che anche se invece di usare 20, uso 18, 19 o 21, lo spettro non cambia di tanto. Quando invece poi si inizia ad avere un ordine ragionevolmente molto più alto di 20, tipo 30, 35, di nuovo lo spettro cambia e compare qualcosa come nel caso 3). Nella pratica questo vuol dire che quando si è nei dintorni dell'ordine corretto, se si aggiunge o toglie una radice, non cambia molto => una cosa che si può fare è quella di guardare come sono fatti gli spettri per qualche ordine basso (5,6,7), poi guardo come sono fatti gli spettri per qualche ordine più alto (18,19,20) e poi per qualche ordine molto alto (30,31,32). Se si nota che da 18 a 20 non cambia niente, ma cambia di più quando gli ordini sono bassi o quando sono alti, evidentemente l'ordine giusto sta tra 18-20. Nessuno discuterà MAI il fatto che è stato scelto un ordine n piuttosto che $n-1$ o $n+1$, cioè non esiste l'ordine corretto reale, ma esiste un modello adeguato al nostro segnale. Quindi se ad esempio, l'ordine corretto teorico fosse stato 20 e invece di 20, scelgo 18 o 21, va bene lo stesso; i problemi, come nel caso 3), iniziano a comparire quando si usano ordini tipo 40, 35, cioè quando si esagera con l'ordine.

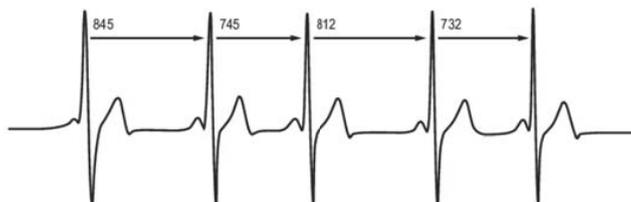
Può capitare ogni tanto che il diagramma della varianza non è facilmente interpretabile come lo facciamo noi, perché il segnale può non derivare da un modello AR e quindi ovviamente bisogna fare attenzione. Questo chiude gli aspetti teorici della stima spettrale parametrica; adesso però bisogna imparare a farla.

Facendo così, trovo che tutte le volta la distanza è leggermente differente. I numeri che sono scritti in figura (845, 745, 812, 732), hanno un'unità di misura che in questo caso è *ms*. Un soggetto sano a riposo ha una frequenza cardiaca di 60 battiti/minuto, ma in realtà non è 60 battiti/minuto, ma 60 è un numero comodo perché all'incirca ogni battito dista dall'altro 1 secondo, che sarebbero 1000 *ms*. In realtà la frequenza cardiaca è un pochino più alta di 60 battiti/minuto e quindi la distanza media è un po' più bassa (circa 850-900 *ms*). Quindi quei numeri sono la misura in *ms* di quanto dista un battito dal successivo.

Se avessi un segnale ECG "bello", cioè ad elevato rapporto segnale/rumore (quindi niente artefatto da movimento, niente tremore muscolare, niente interferenza di rete, ecc.), e dovessi calcolare la distanza che c'è tra un battito e il successivo, sarei capace?

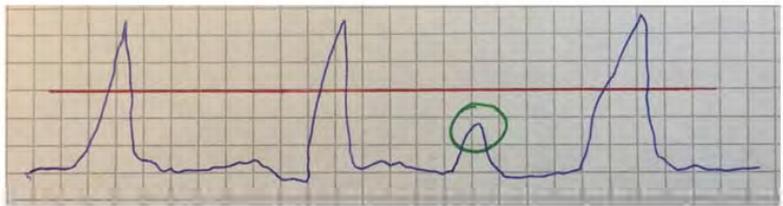
Se devo calcolare la distanza che c'è tra un battito e quello successivo, del complesso delle onde PQRST, devo prendere 2 punti comodi e il punto più comodo è ovviamente l'onda R (perché è quella che o in positivo o in negativo, vedo meglio). Quindi un primo modo potrebbe essere quello di individuare la posizione di tutte le onde R e poi calcolo la distanza che c'è tra un'onda R e la successiva. Il problema a questo punto si riduce a come trovare tutte le onde R, dato un segnale ECG.

Se il segnale ECG è fatto come disegnato qui:



=> non c'è problema, nel senso che basta tirare una bella soglia e poi si va vedere qual è il massimo sopra ogni soglia e trovo automaticamente tutte le onde R. Per trovare il massimo si hanno a disposizione solo due istruzioni su Matlab: *diff* che fa la differenza tra un campione e il successivo e *sign* che è la funzione segno, che vale 1 quando il segno è positivo, -1 quando è negativo → con queste istruzioni, dobbiamo trovare tutti i massimi di un segnale (2 righe: una che usa quelle 2 istruzioni, una con un *if* o un *find*).

Il segnale però non è mai fatto in questo modo, perché le onde R non hanno tutte la stessa altezza, ma hanno un'altezza variabile. Allora succede che si ha questo:



ovvero si hanno onde R di altezze diverse e, tirando una soglia (ROSSA), qualche onda R si può perdere (CERCHIO VERDE).

Se ci venisse chiesto di trovare tutte le posizioni delle onde R, dato un ECG, non sarebbe un problema facile (ci sono molti metodi, alcuni anche complicati); ma il concetto sarebbe quello proprio quello di misurare quanto ogni battito dista dall'altro, basandosi sulla distanza che c'è tra un'onda R e quella successiva. Da qui si crea un segnale di supporto che è fatto in questo modo:

1) la prima è un problema di risoluzione (ed è la principale);

2) la seconda è una frequenza di campionamento non costante (meno importante).

Sulla prima non possiamo ancora dire niente perché non sappiamo come è fatto lo spettro di questo segnale; sulla seconda invece qualcosa si può dire perché questo segnale è campionato ad una frequenza di campionamento non costante, perché si ha un punto ad ogni battito, però ogni battito, temporalmente, dista l'uno dall'altro di una quantità diversa. Quindi se esprimiamo la frequenza di campionamento in battiti, allora la frequenza di campionamento è costante, e vale 1. Però voglio andare a denormalizzare battiti rispetto ad Hz (visto che non si sa mediamente quanto distano i battiti) => se si esprime in Hz come frequenza di campionamento, ovviamente sarà una frequenza di campionamento non costante e quindi certamente questo segnale ha già la seconda caratteristica.

Usiamo questo esempio per fare una domanda che abbiamo tralasciato. Una regola della stima spettrale non parametrica era che se non si levava il valor medio, era un errore grave; per la stima spettrale parametrica non abbiamo mai parlato del valor medio o di rimozione del valor medio.

Questo segnale è ovviamente a valor medio diverso da 0. => è obbligatorio rimuovere il valor medio o no?

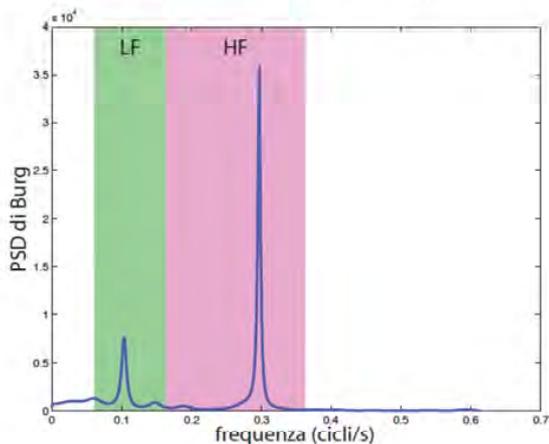
La risposta più facile ed intuitiva è questa: abbiamo visto prima che il valor medio del segnale è, grossomodo, proporzionale alla frequenza cardiaca (è l'inverso della frequenza cardiaca). Siccome si è detto che la frequenza cardiaca non ci interessa (tanto si può calcolare in altri modi), ma ci interessano le oscillazioni, allora il valor medio si può togliere perché tanto non si perde informazione. Questo è corretto, perché se si ha un segnale a valor medio diverso da 0 e si toglie il valor medio, non si sbaglia mai => procedere alla stima spettrale parametrica con un segnale a valor medio nullo, non è un problema.

Però la domanda era diversa: in stima spettrale non parametrica, non rimuovere il valor medio è un errore grave; in stima spettrale parametrica, è ancora un errore grave o no?

Nella stima spettrale tradizionale lasciare il valor medio è un errore molto grave perché c'è il leakage di potenza in continua, perché correlogramma e perdiogramma sono operatori quadratici (nascosto da qualche parte c'è un modulo quadro: per il periodogramma è il modulo quadro della FFT, per il correlogramma è nascosto nella ACS che essendo x per il suo complesso coniugato, vuol dire modulo quadro). Quindi sono due operatori quadratici, e il quadrato genera doppi prodotti che abbinato al leakage di potenza, ogni volta che c'è un'armonica, crea casino per niente. Perciò quello che si può dimostrare è che se si ha un valor medio diverso da zero, si ha una bella componente in continua che si spalma su tutte le altre frequenze. Dato che si è detto che la continua non ci serve, va tolta e quindi si evita la polarizzazione dovuta dalla continua, che inevitabilmente ci sarebbe. Questo è il motivo per cui nella stima spettrale tradizionale il valor medio deve essere tolto obbligatoriamente.

Nella stima spettrale parametrica il problema del leakage di potenza c'è? Non si è parlato né di finestratura, né di consistenza, né di riduzione della variabilità, perché tanta si ha un modello e di quel modello si conosce tutto perché lo abbiamo fatto noi → su questo modello si ha, virtualmente, risoluzione infinita, non ci interessa nulla della polarizzazione e della consistenza, al massimo il modello può essere sbagliato, o modellizzato male il segnale. Allora nella stima spettrale parametrica succede che, venendo meno il problema della polarizzazione e il problema della consistenza, il fatto di lasciare il valor medio non è drammatico, ma si avrà una componente in continua; siccome si è detto che ogni componente in più vuol dire almeno una radice in più (quindi un ordine più alto di uno), si sa già che tra un segnale a valor medio nullo e lo stesso segnale a valor medio diverso da zero, si dovranno usare 2 modelli di ordine diverso => non rimuovere il valor medio quando si fa stima spettrale parametrica, non è un problema, al contrario della stima spettrale non parametrica. Quindi

Slide 46: SPETTRO AR DI BURG (SOGGETTO A RIPOSO)



Lo spettro AR di Burg ha 3 bande di frequenza che sono esattamente le stesse bande di frequenza viste nei segnali di autoregolazione cerebrale. Si era detto che c'era una banda B, una banda M, una banda R e una banda P => sul segnale di variabilità cardiaca, la banda P (quella della pulsatilità cardiaca) non si vede perché è stata già tolta. Infatti avendo fatto la distanza tra un battito e il successivo, l'informazione della banda P, è nel valor medio (che non ci interessa e lo leviamo).

Le altre 3 bande erano:

1) la banda B, a frequenze molto basse (decine di mHz), che era una banda relativa all'autoregolazione metabolico, ritmo sonno-veglia, ecc. Nel segnale di variabilità cardiaca questa banda è presente, ma non si usa dal punto vista clinico => la banda B, che era a frequenze molto basse (tra 20 e 60 mHz), esiste ma non ha particolare importanza.

2) la banda M o LF (bassa frequenza) va da circa 60 a 140 mHz ed la banda che era centrata approssimativamente a 100 mHz. Questa banda è l'impronta del sistema simpatico, cioè è l'impronta di tutti quei sensori di pressione e di composizione ematica del sangue che agiscono sul tono della parete arteriosa.

3) poi c'è una banda a frequenza più alta, che era stata chiamata banda R, e che qui si chiama banda HF (alta frequenza), che è la banda del ritmo respiratorio.

Quindi gli attori che di fatto agiscono sulla variabilità cardiaca, di fatto, sono 3: metabolismo, aspetto vascolare ed aspetto respiratorio; ma i due più importanti sono la parte vascolare e respiratoria. Questo non è un fatto più sorprendente di tanto, perché il cuore batte al suo ritmo, però bisogna anche tener conto delle condizioni che si trova al contorno, cioè: il sangue spinge un fluido incompressibile all'interno di un albero vascolare; se l'albero vascolare è in condizioni di vasodilatazione o in condizione di vasocostrizione, offre una resistenza diversa => questo vuol dire che la pompa cardiaca va un po' più veloce, o un po' più lenta, a seconda di qual è lo stato vascolare (perché fa un po' più fatica o fa un po' meno fatica).

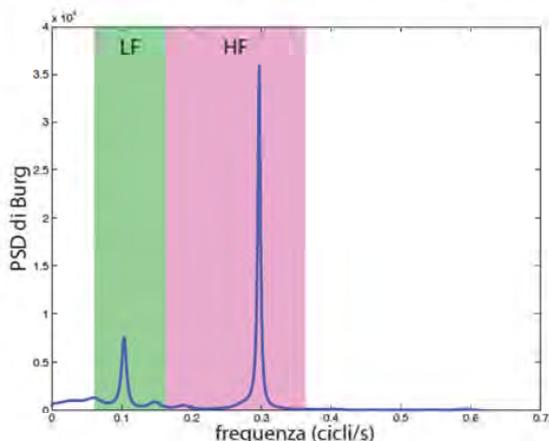
Allo stesso modo per la respirazione, quando respiriamo durante la fase di inspirazione ed espirazione, si varia la pressione intratoracica; variando questa pressione, si ha che la pompa cardiaca, meccanicamente, batte in un ambiente ad alta pressione o a bassa pressione => di nuovo fa più o meno fatica e quindi accelera un po' o rallenta un po'. Per cui è naturale che non si avranno mai battiti equidistanti uno rispetto ad un altro.

Se si vuole far sì che sul segnale di variabilità cardiaca la respirazione non abbia effetto, ovviamente al soggetto bisogna chiedere di sospendere la respirazione (di fare apnea). Ma normalmente più

Allora si può dire che a riposo si ha una certa divisione di potenza; quando si dà uno stimolo, si va a vedere come la ripartizione relativa delle potenze cambia e da qui si prende la decisione per capire se la variabilità è corretta oppure no → in questo caso è corretta. Ci sono però dei soggetti che, una volta che si tirano su in piedi, questo incremento della potenza LF non ce l'hanno: si chiama *scompenso cardiaco* → vuol dire che la variabilità cardiaca c'è, ma non risponde agli stimoli esterni, quindi non si è in grado di compensare variazioni di pressione, gravità, metabolismo, ecc.. (non è una bella cosa).

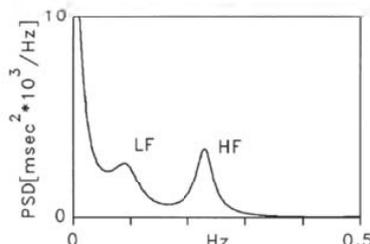
Se prendo il soggetto da seduto e lo metto su una cyclette e si fa pedalare, delle due bande aumenta quella della respirazione e quindi della *torta* si allarga la fetta HF e si riduce la fetta LF. Perciò è anche facile vedere da un metodo di questo tipo se, clinicamente, la variabilità cardiaca è corretta oppure no.

Perché si è detto che questo segnale viene *normalmente* analizzato con stima spettrale parametrica? Certamente non ha una frequenza di campionamento costante, ma poi avendo questo spettro:



si hanno esattamente le stesse problematiche di risoluzione che aveva il segnale di autoregolazione, il segnale di spettroscopia. Cioè, siccome le frequenze sono dell'ordine di centinaia di mHz, una banda attorno a 300 mHz, se si vuole vedere bene allora bisogna avere una risoluzione di qualche mHz → cioè una risoluzione di qualche mHz vuol dire segnali temporalmente lunghissimi: non si può più dire che un segnale biologico su tempi più lunghi di decine di minuti, sia ragionevolmente stazionario. Allora questo è il classico segnale che ha tutte e due le problematiche: problema di risoluzione e problema di frequenza di campionamento non costante. La cosa positiva è che se si riesce a fare bene stima spettrale parametrica, non si acquisisce 100 ore di segnale ECG ma quando il soggetto è sdraiato si acquisisce un po' di ECG, poi si fa tirare su il soggetto e si acquisisce un altro po' di ECG, poi si calcolano i segnali R-R e fine. Ma quanto ECG? Per ottenere questo spettro parametrico AR (l'ordine non lo so, ma per i segnali di variabilità cardiaca gli ordini sono grossomodo tra 10 e 50):

REST



quanti secondi di ECG sono stati acquisiti?? Non si sa, ma se ragioniamo dal punto di vista teorico, non abbiamo mai detto quanto deve essere lungo $x[n]$ per modellizzarlo; ma certamente sappiamo che se il nostro modello ha un ordine pari a 20, per risolvere il sistema lineare dobbiamo avere almeno 20

SESTO LABORATORIO: ANALISI DELLA VARIABILITÀ CARDIACA E SEGNALI CORRELATI

Sono forniti già i segnali RR e non gli ECG, acquisiti da un soggetto in 3 condizioni diverse:

- 1) supina
- 2) in piedi
- 3) durante un esercizio (forse è la pedalata in bici)

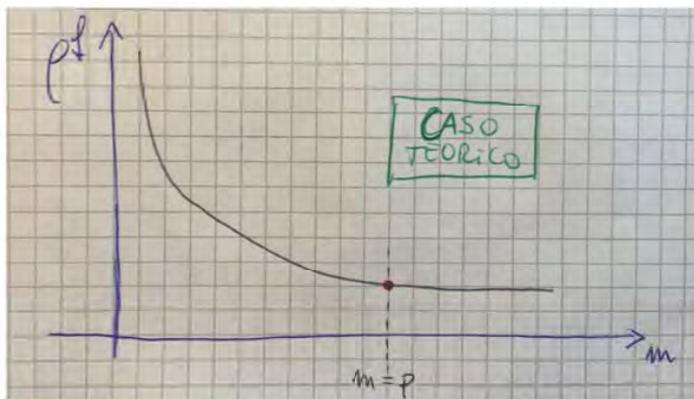
Quello che è stato misurato è il segnale RR (già espresso in ms) e poi un certo numero di segnali di appoggio, che però in questa esercitazione non usiamo (pressione sistolica e flusso respiratorio in litri/secondo).

Normalmente, quando si fanno dal punto di vista clinico studi di questo tipo, difficilmente si misura una cosa sola, ma sempre un paio => oltre a vedere come varia la variabilità cardiaca, si vede anche come varia la pressione (per assicurarsi che non ci sia un problema di tipo vascolare) ed eventualmente si vede anche come varia il ritmo respiratorio, dato che anche quello influisce sulla variabilità cardiaca.

Sul testo sono scritte un po' di cose su come sono stati organizzati i dati, qual è la risoluzione per ogni canale, ecc...e poi le domande sono le seguenti:

1. *imporre un modello di tipo AR, ottenendo la sequenza dei coefficienti mediante il metodo di Burg. **Al fine di stabilire l'ordine del modello**, calcolare la varianza dell'errore per ordini compresi tra 2 e 50. Scegliere come ordine del modello il minimo valore per cui la varianza dell'errore risulta non superiore alla varianza asintotica (cioè quella ottenuta per ordine pari a 50) aumentata del 5%.*

Cosa vuol dire?? Nella pratica vuol dire la seguente cosa: abbiamo visto che la varianza dell'errore (quella che abbiamo chiamato varianza dell'errore di predizione forward in funzione dell'ordine del modello) era una roba che teoricamente doveva essere così:



Cioè c'era un punto, per esempio qua (ROSSO), da cui in avanti la varianza non cambiava più e questo punto era quel punto in cui l'ordine nel predittore lineare era uguale all'ordine del modello AR che aveva generato il segnale, cioè era l'ordine corretto. Allora, per un certo numero di cose che abbiamo visto, questo diagramma della varianza non lo otterremo mai perché ci sono delle approssimazioni numeriche, perché il segnale non è effettivamente derivante da un modello AR e noi tentiamo di modellizzarlo con un modello AR e quindi non ce la facciamo, ecc.. => il diagramma della varianza, normalmente, è più complicato e quindi se quello precedente era il CASO TEORICO, nel CASO PRATICO ci possiamo aspettare che il diagramma della varianza, ad esempio, abbia delle oscillazioni fatte così:

costante in secondi. Allora si fa la seguente cosa: si prende il segnale RR, si calcola il valore medio, questo valor medio è proporzionale alla frequenza cardiaca che ci dice quanti battiti ci sono in 1 minuto e si usa quel valore, per denormalizzare l'asse delle frequenze, in modo da riportarlo non in cicli al battito, ma in cicli al secondo, così almeno si sa dove sono le bande.

3. Ripetere la procedura per tutti e tre i file, osservando come diverse componenti spettrali risultino prevalenti nelle diverse condizioni sperimentali.

4. Ripetere la procedura utilizzando il metodo della covarianza modificata.

5. Ripetere la procedura utilizzando il metodo di Yule-Walker.

VEDIAMO IL CODICE MATLAB:

L'istruzione fondamentale si chiama: `arburg` → calcola i coefficienti e l'errore di predizione (cioè la varianza dell'errore), per un modello AR, di un certo ordine, quando i coefficienti sono calcolati secondo il metodo di Burg (riga 47).

Per il metodo della covarianza modificata, l'istruzione è: `armcov` (riga 65).

Per il metodo di Yule-Walker, l'istruzione è: `aryule` (riga 65).

Il ciclo `for` (riga 46-48, 64-66, 82-84) per un ordine che va da 2 a 50, considerando i parametri che vuole `arburg` (o `armcov` o `aryule`), calcola i parametri e la varianza dell'errore. Poi bisogna calcolare il diagramma della varianza e imponendo una soglia del 5%, si troverà il minimo valore che non supera la varianza asintotica del 5% della soglia impostata => quello sarà l'ordine del modello. Una volta che si ha l'ordine, si può calcolare lo spettro: come esisteva `pwelch`, esiste `pburg` (riga 57), e quindi dato il segnale e dato l'ordine, automaticamente vengono restituiti l'asse delle frequenze e lo spettro in modo che se si fa: `plot(w, Pb)`, disegna esattamente lo spettro AR, con i parametri calcolati secondo il metodo di Burg. Ma quello che si diceva di avere un asse normalizzato, non ci interessa => denormalizziamo in Hz o cicli/secondo che c'è la stessa cosa, tenendo conto che bisogna dividere per quello che è il valor medio del segnale, che utilizziamo come frequenza media, che è la frequenza cardiaca. Allora viene calcolato il valor medio del segnale diviso per il numero di campioni e viene utilizzato al denominatore per denormalizzare l'asse delle frequenze => il plot che si vede, è come se fosse tarato in Hz (come usiamo normalmente).

Nel codice, il valor medio viene rimosso dal segnale, qua : riga 40. Quello che dobbiamo fare è: togliere il valor medio e procedere e poi provare a commentare la riga 40 (senza quindi togliere il valor medio) e vedere cosa succede, iniziando a vedere cosa succede nel diagramma della varianza (cambia?non cambia?cambia l'ordine?bisogna fare scelte diverse?è più difficile?è più facile?) e poi vedere come cambiano gli spettri (Burg, covarianza modificata e Yule-Walker) tra quando si tiene il valor medio e quando si toglie il valor medio.

Inoltre, dato che dal punto di vista clinico, quello che importa è la relazione tra potenza nella banda LF e potenza nella banda HF, bisogna calcolare le potenze. Allora:

6. Nei tre casi, calcolare il rapporto LF/HF dividendo la potenza dello spettro AR nella banda LF (60 - 140 mHz) per quella nella banda HF (140 - 300 mHz). Tenere conto del fatto che gli estremi della banda HF potrebbero variare a seconda dello stato del soggetto e, quindi, del ritmo respiratorio.

Questo punto serve per vedere se effettivamente la banda dominante è quella in cui il soggetto è sdraiato e se è diversa da quella in cui il soggetto è seduto e se è diversa da quella in cui il soggetto effettua un esercizio => si fa qualche considerazione sullo stato del soggetto.

11/05/2015

Time - Frequency Distributions

Questo argomento ci porta un passo avanti rispetto a quanto visto la volta scorsa: non bisogna mai perdere di vista l'ottica fondamentale → il problema di base che abbiamo è quello di vedere come è fatto lo spettro di un segnale. A volte ci riusciamo senza particolari problemi perché il segnale è stazionario in senso lato e lungo a sufficienza => ci possiamo permettere di usare la stima spettrale tradizionale con una risoluzione spettrale teorica che ci va bene e tutto sommato, l'unica cosa a cui bisogna fare attenzione, sono le caratteristiche degli stimatori, che non sono granché. Però una volta che si è capito che chi più e chi meno sono polarizzati, chi più e chi meno sono poco consistenti e si sa come tenere a bada polarizzazione e consistenza, in qualche modo ce la caviamo.

Qualche problema in più lo creano quei segnali che, o sono stazionari per un tempo molto breve, oppure sono essi stessi molto brevi, per cui la stima spettrale tradizionale non funziona, non è applicabile. Allora abbiamo visto che in quel caso ce la possiamo cavare con la stima spettrale parametrica; ma con la stima spettrale parametrica, comunque, in un pezzo lungo o corto che sia, il segnale deve ancora essere stazionario in senso lato, ovvero deve essere ancora assimilabile ad un processo WSS ed ergodico (ergodico non lo diciamo mai, perché si è detto fin dall'inizio che lo diamo per scontato: per noi non c'è modo né di verificarlo, né di verificare il contrario => per noi è ergodico e basta).

Adesso facciamo un passo ulteriore: come si fa con quei segnali che proprio stazionari non lo sono per natura (neanche su finestre cortissime)? Allora, in questo caso, bisogna in qualche modo capire come riuscire a fare a meno del teorema base della stima spettrale, che è il teorema di Wiener-Khinchin → il teorema di W-K dice che il segnale deve essere stazionario in senso lato ed ergodico. Ora, stazionario in senso lato non lo è più, e quindi il teorema di W-K non vale più => non è più vero che la trasformata di Fourier della ACS è lo spettro del segnale e quindi bisogna capire come girarci attorno o come fare in modo che il teorema di W-K non ci complichino troppo la vita.

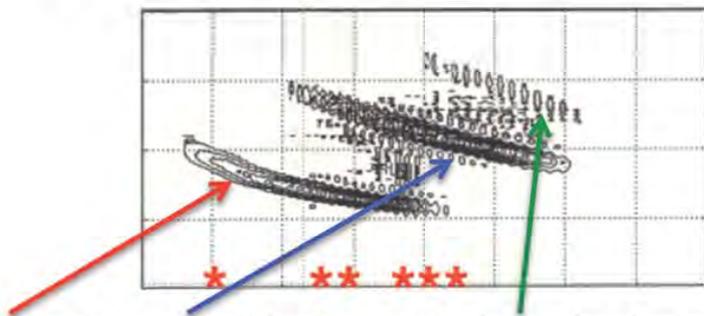
Ce ne sono di segnali biologici che sono non stazionari, e hanno una non stazionarietà molto forte. Un esempio semplice, è il segnale EMG: si è detto che il segnale EMG è stazionario in senso lato se la contrazione è isometrica e a forza costante → se una di queste due condizioni non è soddisfatta, diventa non stazionario. Allora se ad esempio voglio usare la stima spettrale in applicazioni dove la contrazione muscolare è dinamica (medicina del lavoro, medicina dello sport, medicina occupazionale, riabilitazione, ecc.), vuol dire che siamo in un caso in cui il muscolo produce un segnale che non è stazionario, neanche su periodi brevi. Allora quello che vediamo da oggi in avanti è come cercare di effettuare stima spettrale, nel caso in cui l'ipotesi fondamentale di segnale WSS non si può assumere.

Slide 2 - WHY TIME - FREQUENCY TRANSFORMS?

Tutti quelli che partono a descrivere questo problema, partono tutti dallo stesso esempio che è un segnale biologico: la registrazione ad ultrasuoni del segnale, dell'impulso, emesso da un pipistrello. Se uno guarda solo il segnale, capisce poco.

Il pipistrello è un "dispositivo" ad eco-pulsato → vuol dire che emette un impulso ad ultrasuoni, sente gli echi e siccome è cieco, capisce dove sta rispetto al mondo. Ovviamente, se non sente echi, capisce che non ci sono ostacoli e vola tranquillo; quando gli ritorna un eco, fa come il radar: dall'eco che torna, capisce la direzione e dal tempo di volo, capisce la distanza. Allora il pipistrello è una sorta di "radar volante" ad ultrasuoni => questo è il classico esempio di segnale non stazionario nemmeno in senso lato, perché se andiamo a vedere su grafico bidimensionale che riporta in ascissa il tempo (in questo caso in campioni, ma data una frequenza di campionamento, si può convertire in secondi), e sull'asse verticale la frequenza normalizzata tra 0, che è la continua, e la frequenza di Nyquist (ma anche in questo caso con la frequenza di campionamento, si può tarare in Hz), si ha questo grafico:

Però in realtà non è proprio così, perché se si osserva cosa succede rispetto al tempo, si può vedere che non sono due bande, ma sono tre frequenze distinte:



una è questa qua, una è questa e un'altra è questa piccola qua, che ad un certo istante di tempo hanno un valore e man mano che il tempo passa, hanno un valore di frequenza che diminuisce.

Cioè il pipistrello che cosa fa? Emette un impulso che non ha frequenza costante, ma ha una frequenza che varia nel tempo → tutti i volatili fanno così e anche tutti i cetacei, anche il canto dell'usignolo o del canarino non sono segnali monofrequenza, ma invece hanno una frequenza che varia nel tempo. A noi sembra che ci sia una sola frequenza, ma in realtà non è vero perché ad ogni istante di tempo, c'è una frequenza diversa. Allora, se si ha un segnale con una frequenza variabile, potrà mai essere stazionario in senso lato?? Ovviamente no, perché stazionario in senso lato vuol dire che il contenuto spettrale, cioè la funzione di autocorrelazione, non varia nel tempo (dipende dal ritardo, ma non dipende dal tempo) => certamente se cambia la frequenza, certamente dipende dal tempo.

Quindi un segnale a frequenza variabile, non potrà mai essere stazionario in senso lato.

Questo segnale però, è ancora più complicato, perché si è detto che non ci sono solo due bande, ma ci sono tre frequenze e ognuna delle quali ha una modulazione di frequenza decrescente; quindi il canto del pipistrello è un segnale tri-componente, dove ogni componente ha una modulazione quasi lineare di frequenza (modulazione in frequenza vuol dire che quella frequenza diminuisce nel tempo).

Però se si legge attentamente questo grafico, in funzione del tempo, si scopre che ad esempio, a questo istante *, all'interno del segnale, è presente solo una componente frequenziale; ad esempio a quest'altro istante **, ne sono presenti due; a questo istante *** ce ne sono presenti tre. Quindi il pipistrello emette un impulso un po' particolare, che è fatto da tre componenti che si innestano a tempi diversi, a frequenza diversa, tutte e due decrescenti nel tempo => è un segnale complicato.

Perché il pipistrello si complica così la vita? Noi la vediamo dal punto di vista ingegneristico, non biologico => il pipistrello sa bene che, più gli ultrasuoni hanno frequenza elevata, meno si propagano (ovvero più è alta la frequenza degli impulsi, più difficile è la propagazione, perché il mezzo assorbe di più, anche l'aria). Allora il pipistrello sa che le frequenze molto alte si propagano poco e allora le emette più tardi; le frequenze più basse, invece, si propagano più lontano e siccome bisogna dare il tempo all'impulso di arrivare più lontano e di tornare indietro, logicamente le deve emettere prima di tutte le altre. Quindi con questa componente (FRECCIA ROSSA), il pipistrello fa una scansione degli ostacoli che ci sono in lontananza; con questa frequenza (FRECCIA BLU) fa una scansione degli ostacoli che ci sono a metà; con questa frequenza (FRECCIA VERDE) fa una scansione degli ostacoli vicini. Il pipistrello, inoltre, sa anche che più alta è la frequenza, maggiore è la risoluzione con cui trova un oggetto => il pipistrello si fa una mappa all'incirca di quelli che sono gli ostacoli lontani, più o meno accurata di quelli che sono gli ostacoli a metà, molto accurata di quelli che sono gli ostacoli vicini. Quindi tre frequenze diverse perché la propagazione è diversa a seconda della frequenza. Inoltre gli dà una modulazione di frequenza perché è l'unico modo per sapere che quell'eco è effettivamente quello relativo a quella frequenza, cioè siccome sa che il mondo è disturbato ed arrivano suoni di tutti i tipi, per sapere che quello è effettivamente l'eco proveniente dal suo impulso, il pipistrello gli dà una

Quindi il problema che avremo sarà quello di avere un segnale con un contenuto spettrale che in funzione del tempo cambia, e noi dobbiamo rincorrerlo, senza violare W-K (dobbiamo fare in modo che W-K non sia violato → se lo violiamo, otteniamo di sicuro un risultato sbagliato).

Slide 4-5-6 - WIENER-KINTCHINE THEOREM

$$P_{xx}(f) = \int_{-\infty}^{+\infty} r_{xx}(\tau) e^{-j2\pi f\tau} d\tau$$

Questo è il teorema di W-K: lo spettro del segnale è la trasformata di Fourier della ACS, sotto l'ipotesi che il segnale sia un processo casuale, stazionario almeno in senso lato ed ergodico.

Useremo questa notazione che ci dice che lo spettro (P_{xx}) è la trasformata di Fourier dalla variabile τ alla variabile f , della sequenza di autocorrelazione (che ovviamente è in funzione di τ). Allora si scrive a pedice di F (che simboleggia la trasformata di Fourier) la variabile di origine, che in questo caso è il ritardo, e la variabile di trasformazione, che in questo caso è la frequenza:

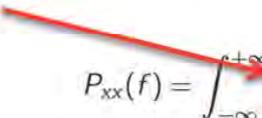
$$F_{\tau \rightarrow f} \{r_{xx}(\tau)\}$$

Se con la stessa notazione avessimo scritto la trasformata di Fourier di un segnale, avremmo scritto che: $X(F) = F_{t \rightarrow f} \{x(t)\}$, per dire che t è la variabile originale mentre f è la variabile di trasformazione.

In questo caso sembra banale, perché la funzione di autocorrelazione dipende solo dal ritardo e ovviamente voglio trasformare il ritardo in frequenza; però, quando si avrà un piano bidimensionale, si avranno funzioni che dipendono da 2 variabili e allora bisogna capire quale trasformiamo in che cosa (=> questa notazione si usa per non scrivere ogni volta integrali molto lunghi).

Sotto le ipotesi dette, il teorema di W-K ci assicura che lo spettro è reale, simmetrico, definito positivo e sempre uguale, cioè vuol dire che: se ad esempio si ha un segnale lungo 100 secondi che è stazionario in senso lato => si può prendere un pezzo qualsiasi di segnale, lungo quanto ci pare, ma tanto il risultato che si ottiene è sempre lo stesso. Ovviamente più lungo si prende, più si ha risoluzione spettrale, più è corto, meno risoluzione si ha, ma lo spettro è sempre lo stesso.

Il fulcro ovviamente sta qua:

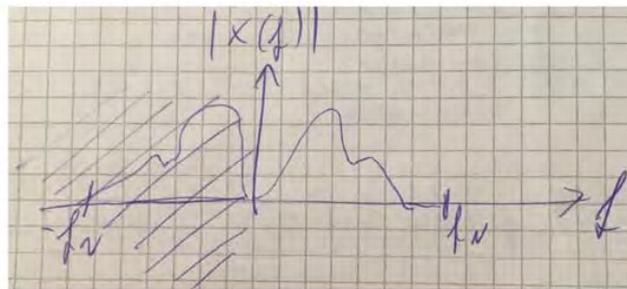
$$P_{xx}(f) = \int_{-\infty}^{+\infty} r_{xx}(\tau) e^{-j2\pi f\tau} d\tau$$


Questa quantità r_{xx} , ovvero la funzione di autocorrelazione, non dipende dal tempo ma solo dal ritardo => dato che la variabile tempo non esiste, allora quindi questo operatore non dipende dal tempo (una volta selezionato il pezzo di segnale, quello è e basta).

Questo approccio abbiamo detto che non va bene per certi segnali, perché non è adatto ad esprimere effettivamente come è fatto il segnale all'interno.

Slide 7 - WHAT IF THE SIGNAL IS NOT WSS

Qui c'è un segnale un po' particolare che in realtà tenta di modellizzare un po' il segnale visto prima, che è il segnale che si chiama *CHIRP* → si chiama così perché gli inglesi il canto dell'uccellino lo chiamano chirp: un segnale che ha un contenuto di frequenza modulata che fa finta di simulare il canto di un uccello.



Con f_N = frequenza di Nyquist e ovviamente la trasformata di Fourier occupa anche il semiasse negativo perché è simmetrica, per esempio in questo caso, rispetto all'origine.

Allora il segnale $x(t)$ è un segnale reale e come si sa, un segnale reale ha una trasformata di Fourier che è quasi sempre complessa, a meno di un caso particolare \rightarrow ovvero che il segnale $x(t)$ sia pari: in questo caso la trasformata di Fourier è reale, ma questo è l'unico caso, perché normalmente la trasformata è una quantità complessa.

Questo segnale disegnato, è un segnale scritto non in forma analitica, ma è un segnale normale, un segnale temporale; se vogliamo scrivere questo segnale in forma analitica, cosa dobbiamo fare?? Scrivere un segnale in forma analitica, vuol dire metterlo in una forma per cui la sua trasformata di Fourier è priva della parte del semiasse negativo, cioè è senza la replica spettrale (noi in genere la ignoriamo a livello di rappresentazione, ma a livello di calcolo ne tengo in considerazione, infatti se metto NFFT=256, di punti utili sul semiasse di interesse, ne ho solo la metà +1 perché l'altra metà è andata alla sua replica spettrale \Rightarrow si calcola, anche se non si rappresenta).

\Rightarrow vogliamo un segnale dove il semiasse negativo è 0! Per far questo, bisogna trasformare il segnale in quella che è la sua versione analitica: $x_a(t)$. Il segnale analitico associato ad un segnale reale $x(t)$, non è altro che il segnale stesso + una parte immaginaria che moltiplica \tilde{x} , che è la trasformata di Hilbert di $x(t)$:

$$x_a = x(t) + j\tilde{x}(t)$$

La trasformata di Hilbert si chiama anche trasformata coseno, cioè vuol dire che è identica alla trasformata di Fourier, solo che invece di moltiplicare per $e^{-j\omega t}$, si moltiplica per $\cos(2\pi f t)$, cioè ci dimentichiamo la parte immaginaria.

\Rightarrow la trasformata di Hilbert è uguale alla trasformata di Fourier, non cambia niente: integrale del segnale, solo che invece di essere moltiplicato per l'esponenziale di Fourier, è moltiplicato solo per la parte reale dell'esponenziale di Fourier, che è il coseno.

Se prendo un segnale particolare, per esempio:

$$x(t) = e^{j2\pi f_0 t} = \cos(2\pi f_0 t) + j \sin(2\pi f_0 t)$$

cioè si scopre che l'esponenziale di Fourier non è altro che il segnale analitico di una sinusoidale (ovvero una sinusoidale in forma analitica): la parte reale è un coseno, la parte immaginaria è la trasformata di Hilbert del coseno. Infatti, se si fa la trasformata di Fourier di $e^{j2\pi f_0 t}$, non si trovano 2 δ , ma una sola δ a frequenza f_0 (la parte di semiasse negativo è nullo).

In questo modo, avendo un segnale complesso, $x_a = x(t) + j\tilde{x}(t)$, la sua trasformata di Fourier non è più necessario che sia simmetrica (infatti non lo è perché c'è la parte immaginaria che cancella il semiasse negativo).

Questa introduzione sul segnale analitico ci serve perché si vedrà che per fare due calcoli quando introduciamo questi strumenti matematici tempo-frequenza, che sono matematicamente più

dove:

$$- \varphi(t) = \pi \alpha t^2$$

- α è il coefficiente di modulazione di frequenza \rightarrow da questo si può capire, ad esempio, quanto varia la frequenza in un determinato lasso di tempo.

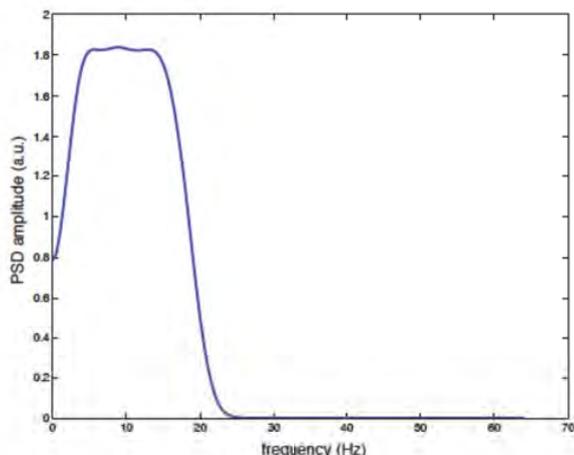
Allora, fermo restando che adesso si sa com'è definito un segnale chirp, in Matlab possiamo sintetizzare un segnale chirp, con una frequenza che in 1 secondo parte da 30Hz e arriva a 60Hz. Si tratta solo di prendere quella retta del grafico precedente, spostarla dandogli un offset, che parte da 30Hz, una pendenza, che dopo un secondo arriva a 60Hz, fare l'integrale e mettere i numeri giusti. Potremmo avere anche, ad esempio, da calcolare un segnale chirp che parte con una frequenza di 60Hz e dopo 2 secondi ha una frequenza di 10Hz (cioè modulazioni decrescente) \Rightarrow si avrà semplicemente un valore di α negativo.

\Rightarrow dobbiamo essere in grado di riuscire a sintetizzare un segnale chirp in Matlab (che non è difficile).

NB: l'asse dei tempi va elevato al quadrato (si ha un t^2).

Questo segnale è quello che abbiamo detto che ha una frequenza che parte da un certo valore (in questo esempio dall'origine) e dopo un po', nel tempo, cresce o diminuisce (in questo esempio è crescente nel tempo).

Se si prova a sintetizzare questo segnale in Matlab e se ne fa il periodogramma o il correlogramma e si rappresenta, si ottiene questa roba qui, che non dice assolutamente come varia in realtà la frequenza istantanea, cioè non si ha idea se si guarda solo lo spettro \Rightarrow è uno strumento inadeguato per seguire quelle che sono le evoluzioni del segnale.



Quindi, la prima idea che è venuta in mente per riuscire a tenere conto di questo problema, è stata quella di usare l'approccio più banale: il segnale non è stazionario nemmeno in senso lato, ma se lo si osserva su una finestra veramente molto piccola, a quel punto è stazionario. Però, così facendo, la risoluzione spettrale diminuisce sempre di più (ma questo si sa già). Questo approccio è un approccio che sacrifica totalmente la risoluzione spettrale, cioè la prima idea è stata quella di dire: "ho un segnale non stazionario, non so bene come trattarlo e per non violare il teorema di W-K riduco talmente tanto la finestra di osservazione, che faccio finta che in una finestra così piccola, il segnale torni ad essere WSS, sacrificando però la risoluzione spettrale".

l'uscita è la somma dell'effetto che questo sistema ha, su ogni singola componente dell'ingresso, senza aggiungere niente e senza togliere niente.

- ⇒ se invece di mettere $x(t')$, metto un segnale che è fatto da 2 componenti: $x_1(t') + x_2(t')$, è vero che la STFT di x è uguale alla somma delle due STFT di x_1 e x_2 ? Sì, perché se ci metto una somma di segnali, separo i due integrali e automaticamente uno è la STFT di x_1 e l'altro è la STFT di x_2 .

Quindi il metodo classico per verificare la linearità di un qualunque sistema è verificare se vale la sovrapposizione degli effetti.

Domanda n°2: la STFT è invertibile? Cioè, se si ha la STFT già fatta, che è una funzione bidimensionale e si ha anche la finestra γ con cui si è ottenuta, riusciamo a ricavare il segnale? Sì => è invertibile.

Quando dobbiamo lavorare con la STFT, uno dei problemi è che abbiamo introdotto questo operatore per fare stima spettrale di un segnale che è non stazionario, e si è detto che il teorema di W-K non lo violiamo perché questa finestra è piccola a sufficienza da poter considerare stazionario un segnale all'interno di quella finestra. Se però abbiamo un segnale che non conosciamo, si sa a priori quanto deve essere lunga la finestra, per assicurare che all'interno della finestra il segnale sia stazionario? No, non c'è metodo => dove cade la STFT?

- 1) sul fatto che stringendo la finestra, si sacrifica inevitabilmente la risoluzione spettrale;
- 2) operativamente, per un segnale di cui non si conosce nulla, trovare la dimensione corretta della finestra, è praticamente impossibile.

=> non c'è la certezza di non violare W-K, prendendo una finestra un po' troppo grande e quindi è un problema di difficile soluzione.

Altra considerazione: noi introduciamo questo strumento per fare stima spettrale, ma la quantità che otteniamo, ha le caratteristiche di uno stimatore di densità spettrale di potenza?

- È reale? No, è complessa perché è una trasformata di Fourier (solo nel caso di segnale pari, è reale) → non ci piace tanto come strumento.
 - È simmetrica? Certo, perché è una trasformata di Fourier.
 - È positiva? Se non è reale, non ha senso porsi il problema che sia positiva o negativa, è un numero complesso.
- ⇒ la STFT non ha nessuna delle caratteristiche che vogliamo per essere uno stimatore spettrale; è solo una trasformata di Fourier, schiacciata nel tempo.

Slide 10 - SPECTROGRAM

Se vogliamo qualcosa simile ad uno stimatore spettrale, non va bene la STFT ma dobbiamo prendere qualcosa che stia alla PSD, come la trasformata di Fourier sta al periodogramma. Tra la trasformata di Fourier e il periodogramma c'è l'operatore *modulo quadro* => se si vuole uno stimatore spettrale, bisogna prendere il modulo quadro della STFT; questa quantità si chiama: SPETTROGRAMMA.

$$SPEC_x^\gamma(t, f) = |STFT_x^\gamma(t, f)|^2$$

Lo spettrogramma è una quantità tempo-frequenza, che viene derivata dalla STFT di un segnale, data una certa finestatura γ , di cui si prende il modulo quadro => in questo modo non si hanno problemi, perché è certamente reale, simmetrica e definita positiva (come il periodogramma, non si hanno problemi); ma non è più lineare, come non lo era il periodogramma.

Quindi si fanno le stesse considerazioni del periodogramma: si devono togliere il valor medio, il trend, ecc..

LABORATORIO 7

La STFT sarà oggetto del laboratorio 7, in cui non si avrà il codice guida => avremo solo delle condizioni ma non la guida.

1. Realizzare una funzione Matlab che calcoli la STFT di un generico segnale come da definizione. Questa funzione deve avere come parametri di ingresso:

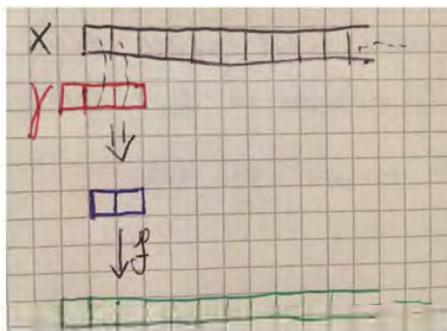
- il vettore contenente il segnale di ingresso;
- il numero di punti su cui calcolare la FFT (indicato generalmente con NFFT);
- la frequenza di campionamento del segnale;
- il numero di punti su cui è definita la finestra di analisi;
- il tipo di finestra.

In uscita la funzione deve restituire la STFT del segnale (che sarà una matrice) ed i due vettori contenenti gli istanti temporali (asse tempi) ed i valori di frequenza (asse frequenze).

Vuol dire che la routine dovrà avere in ingresso: il segnale, la sua frequenza di campionamento, il numero di punti su cui si vuole esprimere l'asse verticale della frequenza e il tipo di finestra che si vuole usare per calcolare la STFT. Quando si dà in ingresso la finestra, il codice deve essere sufficientemente astuto che se come ingresso gli do Hamming 128, lui prende il segnale lungo quanto è, e ci fa passare sopra automaticamente una finestra da 128 campioni; un po' come faceva *pwelch*, che quando si dava il segnale, lungo quel che fosse, e una finestra di una certa dimensione, lui automaticamente spezza il segnale in pezzi di quella dimensione, li finestra, fa il periodogramma semplice e fa la media.

In questo caso bisogna tener conto che non bisogna prendere il segnale e tagliarlo a pezzi, ma bisogna prendere il segnale, iniziare dall'inizio e poi far scorrere la finestra sul segnale, campione per campione. Cioè, quello che dobbiamo fare, è una roba del genere:

x è il nostro segnale che diamo in ingresso, poi abbiamo la finestra γ che scorre sul segnale campione per campione => vuol dire che all'inizio la finestra deve essere centrata sul primo campione del segnale → per essere centrata sul primo campione del segnale, la finestra è centrata così (ROSSO) (quindi bisognerà fare zero padding, mettere degli zeri). A questo punto si fanno i prodotti tra gli elementi del segnale e gli elementi della finestra, e il prodotto risultante, in questo caso, è solo un vettore di 2 campioni che contiene questi prodotti (perché l'altro è nullo) e di questa roba qui (BLU), bisogna fare la trasformata di Fourier.



La trasformata di Fourier, indipendentemente da quanto è lungo il campione di ingresso, ha una lunghezza che è definita da NFFT => il fatto che più campioni siano 0 o meno, non ci interessa. Infatti, siccome io la posso esprimere su quanti campioni voglio, la trasformata di Fourier sarà sempre lunga uguale. Quindi questa (VERDE) è la trasformata di Fourier relativa al primo campione.

15/05/2015

riepilogo..

Riprendiamo dalla definizione di STFT:

$$STFT_x^\gamma(t, f) = \int_{-\infty}^{+\infty} x(t')\gamma(t' - t)e^{-j2\pi ft'} dt'$$

Si era detto che la STFT di un segnale x , calcolata tramite una finestra γ , è una funzione che dipende dal tempo e dalla frequenza ed è un integrale di convoluzione => una finestra γ viene fatta scorrere sul segnale e poi si fa la trasformata di Fourier.

Allora, l'argomento del **laboratorio 7** è quello di scrivere una funzione Matlab che implementi questa roba qua, senza codice guida.

Dobbiamo scrivere una funzione: `stft` che in ingresso prende il segnale (ad esempio x), il numero di punti su cui calcolare la FFT ($NFFT$), la frequenza di campionamento (fc), e poi da qui aprono 2 strade:

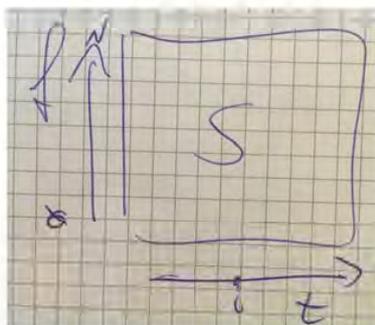
- 1) si può mettere o il numero di punti su cui è definita la finestra di analisi (N_w) e il tipo di finestra (ad esempio `'hamming'`) => all'interno del codice bisognerà generare una finestra (dello stesso tipo specificato) con quel numero di punti;
 - 2) oppure passo alla funzione direttamente la finestra, cioè il vettore $W \rightarrow$ conterrà la finestra e quindi la lunghezza del vettore W è già il numero di punti.
- ⇒ vanno bene entrambe le strade.

In uscita avrò la matrice in cui è contenuta la STFT (ad esempio la chiamiamo S) → una matrice perché essendo una variabile che dipende dal tempo e dalla frequenza, ovviamente abbiamo bisogno di righe e di colonne. Quindi dovrà restituire anche i due vettori contenenti gli istanti temporali (cioè l'asse dei tempi) e valori in frequenza (cioè l'asse delle frequenze), ovvero rispettivamente i vettori t ed f . Allora:

```
[S, t, f] = stft (x, NFFT, fc, Nw, 'hamming') || [S, t, f] = stft (x, NFFT, fc, W)
```

Ora bisogna capire concettualmente come strutturare un programma per fare la STFT: esistono mille modi, questo è quello che forse è più comodo, ma possiamo farlo come ci pare. Allora il consiglio è:

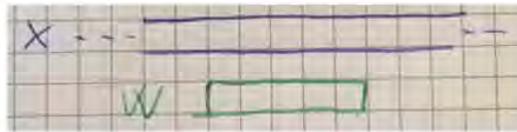
far sì che S sia una matrice dove il tempo è per colonne e la frequenza per righe:



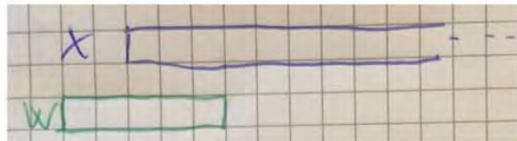
che si metta la continua in basso e la frequenza di Nyquist in alto, oppure viceversa, è irrilevante. L'importante è che ogni riga deve corrispondere ad un valore di frequenza diverso e ogni colonna ad un istante temporale diverso.

Allora, concettualmente, il codice deve prendere la finestra specificata in ingresso, farla scorrere sul segnale e ogni volta che la finestra è centrata su un campione del segnale, bisogna fare il prodotto tra gli elementi della finestra e gli elementi del segnale e di questo prodotto, fare la FFT.

Qui c'è la prima criticità: quando siamo in un caso come prima, cioè la finestra è questa qua, non ci sono problemi.



Quando però la finestra è centrata all'inizio del segnale, bisogna fare attenzione perché non ho elementi a sufficienza dalla parte sinistra:

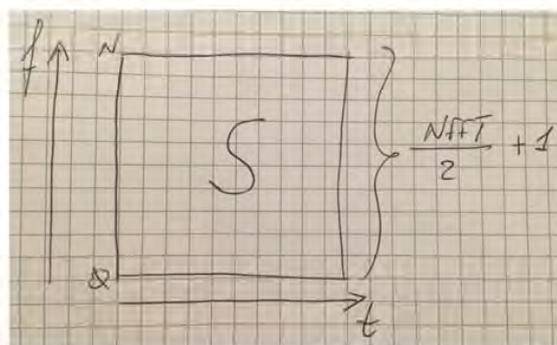


Questo vuol dire che non basta prendere la finestra, farlo scorrere per i che va da 1 a 100 e fare il prodotto; perché quando si è all'inizio del segnale e quando si è alla fine, non si hanno abbastanza campioni (la finestra pende fuori) => bisogna fare zero padding, cioè bisogna gestire questo fatto che finché non ci sono elementi, sufficienti, non possiamo fare il prodotto tra gli elementi del segnale e gli elementi della finestra.

- 2) È chiaro il numero di colonne della matrice S , ma il numero di righe?? È pari a $NFFT$, però siccome quello che andiamo a mettere per colonne è una trasformata di Fourier, quella trasformata di Fourier contiene sia l'asse positivo delle frequenze che l'asse negativo; ma a noi l'asse negativo non interessa, perché è simmetrica (ci dà la stessa informazione) => prima di mettere la trasformata di Fourier all'interno della matrice, tronchiamo la replica spettrale. Per cui il numero di righe vale:

$$\frac{NFFT}{2} + 1$$

- ⇒ vogliamo una matrice S , dove l'asse delle frequenze contiene solo il semiasse positivo delle frequenze e il restante lo tronchiamo, tanto non ci serve. Quindi se prendiamo i primi $\frac{NFFT}{2} + 1$ campioni, si avrà la continua in alto e la frequenza di Nyquist in basso; se si prendono i secondi $\frac{NFFT}{2} + 1$ campioni, il contrario (è giusto in tutti e due i modi).



Queste sono le due maggiori criticità. Se adesso dobbiamo fare una routine Matlab che fa questa roba, l'istruzione che comanda tutto è un enorme ciclo `for` che ciclerà su tutti i campioni del segnale e all'interno di questo ciclo `for` bisogna fare il prodotto, campione per campione, tra il pezzo di segnale corretto (man mano che mi sposto) e la finestra. Di questa parte risultante bisogna fare la FFT con il valore di $NFFT$ corretto, e poi bisogna mettere questo vettore risultante al posto giusto all'interno della matrice S .

```
f = [0:size(S,1)-1]*fs/NFFT           %asse delle frequenze
```

%in questo caso partiamo da 0 e non da 1, perché altrimenti ci perdiamo la continua.

NB: *S* è una matrice complessa, perché contiene metà trasformata di Fourier

ATTENZIONE: In Matlab può essere molto comodo creare matrici e vettori incrementali, del tipo: ho un vettore *a* che contiene un certo numero di elemento e a un certo punto voglio aggiungere un elemento al vettore *a* => aggiungo un elemento *k-esimo* al vettore *a*:

```
a=[ . . . . . ]
```

```
a=[ a k ]
```

Se questa roba è all'interno di un ciclo `for` che va da 1 a 100, il vettore *a* cresce di 100 elementi (1 per ogni ciclo).

Questa cosa è molto semplice, ma crea solo un problema: se devo far crescere un vettore lineare, non ci sono problemi → esempio pratico: ho un tracciato ECG e devo mettere in un vettore l'indice a cui c'è ogni onda R di tutto il tracciato ECG. Allora con un algoritmo vado a trovare la posizione delle onde R e poi avrò un vettore che contiene gli istanti in cui avviene ogni onda R => posso fare una cosa iterativa, partendo dall'inizio: trovo la prima onda R e la metto nel vettore, poi al passo successivo trovo la seconda onda R e la aggiungo al vettore e così via.. e magari all'inizio il vettore non potevo nemmeno inzializzarlo perché non so quante onde R ci sono all'interno di quel vettore. Quindi uso questa strategia perché il vantaggio di Matlab è che non ho bisogno di pre-inizializzare un matrice con delle dimensioni fisse, ma ci metto dentro quello che voglio. Il problema però nasce quando però invece di un vettore, bisogna aggiungere righe, colonne o pezzi ad una matrice, perché si rischia di rallentare in modo tremendo Matlab → infatti Matlab prende la variabile iniziale che ha in memoria, ce ne aggiunge un pezzo, la copia in una variabile di appoggio, cancella la variabile iniziale, riassegna lo stesso nome alla variabile di appoggio => fintanto che si tratta di un vettore, non ci sono problemi; quando però si obbliga Matlab a spostare in memoria matrici di numeri complessi da 1024x1024, ci si accorge che quando si iniziano ad aggiungere colonne, va molto lento perché sta trafficando con la memoria. Matlab tra l'altro è molto conservativo da questo punto di vista, nel senso che fa sempre la stessa cosa: se abbiamo un PC con 4GB di memoria oppure con 8GB oppure con 16GB, li occupa sempre tutti, per cui dopo un po' inevitabilmente rallenta. Allora la pre-inizializzazione delle variabili (specie delle matrici che possono contenere numeri complessi ed essere grandi), è fondamentale per evitare di perdere tempo quando si fanno cicli di questo tipo; ovviamente si può pre-inizializzare tutte le volte che si conoscono le dimensioni.

Poi l'esercitazione dice:

2. Generare una sinusoide in forma analitica della durata di 1 s e di frequenza pari a 20 Hz, campionata a 128 Hz. Verificare il corretto funzionamento della routine implementata al punto precedente. Si consiglia di verificare il risultato graficamente, utilizzando (per la rappresentazione della STFT del segnale) le istruzioni `contour` e `mesh` di Matlab.

Allora, quello che abbiamo visto la volta scorsa sul segnale analitico e quello che ci servirà oggi e quello che vedremo da qui in avanti, è obbligatorio perché per riuscire a capire come funzionano gli strumenti, con le sinusoidi è facile per il motivo che sappiamo già che la rappresentazione tempo-frequenza di una sinusoide è una riga alla frequenza della sinusoide.