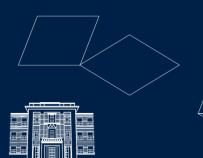
Introduzione alle tecniche di ricostruzione iterativa per la tomografia ad emissione

Focus principale sugli algoritmi ML-EM e OS-EM



Ph.D. Student Dip. Ingegneria dell'informazione Università di Pisa

Corso di Immagini Biomediche, 28 Ottobre 2016







Università di Pisa

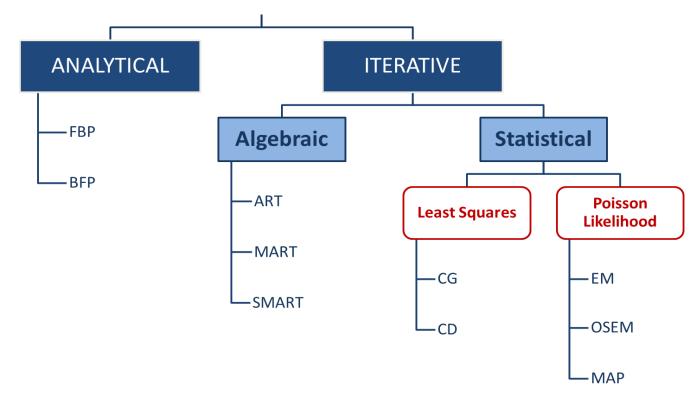


Outline

- > Introduzione
- > Ricostruzione tomografica: un problema lineare inverso
- > Elementi di un algoritmo iterativo
- Criteri di ottimizzazione
- Algoritmi di ricostruzione iterativa
- > Esercitazione



Reconstruction Methods





Metodi analitici vs Metodi iterativi

Metodi analitici (FBP)

Problema matematico basato sull'**inversione della trasformata di Radon discreta**. I dati sono integrali di linea e non si tenta di modellare esplicitamente la randomicità tipica del processo di conteggio dei fotoni .

Metodi iterativi

Generico modello lineare che consente una descrizione dettagliata dei meccanismi di blurring ed attenuazione. Oltre a questo, le tecniche statistiche di ricostruzione cercano di incorporare nel modello anche una descrizione probabilistica del rumore.



Metodi analitici vs Metodi iterativi

Trade-off

Introduzione

Il trade-off tra tecniche iterative ed FBP è quindi una scelta tra accuratezza ed efficienza della ricostruzione.

Tecniche analitiche

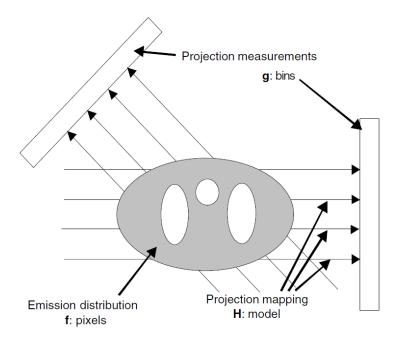
- Modello non fedele alla natura dei dati reali
- Immagini ottenute velocemente ma poco accurate
- > Immagini molto rumorose

Tecniche iterative/statistiche

- Modello matematico molto più complesso
- Non esiste soluzione analitica in forma chiusa
- > Tempi di calcolo molto maggiori



Modello lineare

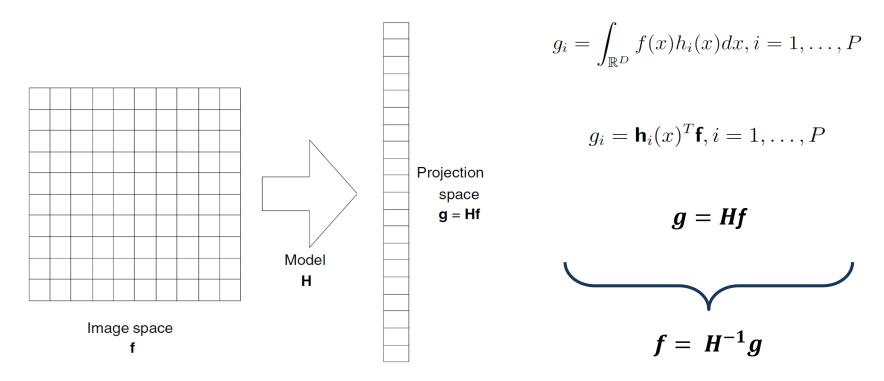


- **g** è l'insieme di proiezioni angolari misurate, il sinogramma
- H è la matrice di sistema, e descrive l'intero processo di registrazione delle immagini. Ogni elemento di H rappresenta il contributo medio di un pixel dell'oggetto al valore di unbin del sinogramma.
- Ogni elemento di f rappresenta un pixel /voxel dello spazio immagine (indifferentemente una fetta 2D o un volume 3D)



Modello lineare

Introduzione



H è grandissima, non quadrata e non invertibile



Modello statistico

Metodi algebrici: interessati al comportamento medio del sistema

Metodi statistici: le misure g_i sono trattate come variabili aleatorie descritte da

distribuzione di probabilità di Poisson (distribuzione tipica di

fenomeni di conteggio)

Una riformulazione più corretta della dovrebbe infatti essere:

$$E[\boldsymbol{g}] = \bar{g}_i = \sum_{i} h_{ij} f_j = \boldsymbol{H}_i \boldsymbol{f}$$

$$p(\boldsymbol{g}|\boldsymbol{f}) = \prod_{i} p(g_i|\boldsymbol{f}) = \prod_{i} \frac{e^{-\boldsymbol{H}_i \boldsymbol{f}} (\boldsymbol{H}_i \boldsymbol{f})^{g_i}}{g_i!}$$

Valore atteso (medio)

Probabilità di misurare g, data un'immagine vera f



Modello statistico

Nota bene

Il modello di Poisson descrive bene i dati di emissione non corretti ed e il più utilizzato nel campo della tomografia ad emissione (PET e SPECT).

Se il sistema di imaging corregge internamente i sinogrammi per le coincidenze random (sottraendo una stima degli stessi dai conteggi totali), la statistica muta.

Solitamente si dovrebbe ricorrere ad un modello **Shifted Poisson**, ma più generalmente si opta per un **modello Gaussiano** più semplice

se il numero medio di eventi e sufficientemente ampio, la distribuzione di Poisson può essere ben approssimata da una gaussiana con media e varianza uguali tra loro e di valore \bar{g}_i



Ogni metodo di ricostruzione iterativo si compone necessariamente di due elementi fondamentali:

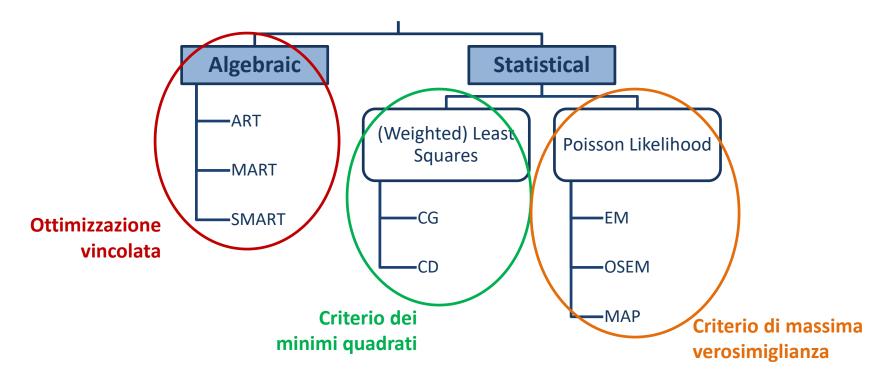
- criterio di ottimizzazione: e il criterio rispetto al quale e possibile determinare quale immagine deve essere considerata la stima migliore dell'immagine vera;
- algoritmo di ottimizzazione: tecnica computazionale finalizzata a cercare la soluzione richiesta dal criterio di ottimizzazione.

Detto in breve ...

Il criterio e la strategia di ricostruzione, l'algoritmo la definizione dei singoli passi necessari ad implementare tale strategia.



ITERATIVE METHODS





Ottimizzazione vincolata (ART)

Criterio di ottimizzazione vincolata

Ricerca di un'immagine che soddisfi una serie di vincoli imposti dai dati misurati e da alcune ipotesi a priori (ad esempio la non-negativita dei pixel). Questa via ha portato alla definizione di una serie di algoritmi che ricadono nella categoria delle tecniche algebriche di ricostruzione (ART).

Il **punto debole** di questi approcci e che non offrono nessun meccanismo che consenta di incorporare un modello statistico esplicito dei dati con cui abbiamo a che fare.



Criterio dei minimi quadrati (LS, WLS)

Criterio LS

Scegliere il valore di f che, se osservato attraverso la matrice di sistema H, garantirebbe delle proiezioni Hf piu possibile simili alle proiezioni osservate g (distanza Euclidea).

$$\hat{\mathbf{f}} = \underset{f}{\operatorname{arg min}} (\mathbf{g} - \mathbf{H}\mathbf{f})^T \mathbf{D} (\mathbf{g} - \mathbf{H}\mathbf{f})$$

$$= \underset{f}{\operatorname{arg min}} \sum_{i=1}^p d_i (g_i - \sum_{j=1}^N h_{ij} f_j)^2$$

D è una matrice diagonale i cui elementi d_i sono $(var[g_i])^{-1}$. Nella tomografia ad emissione, i cui dati sono distribuiti secondo Poisson, la varianza e uguale alla media, per cui $d_i = g_i^{-1}$

$$\hat{\mathbf{f}} = (\mathbf{H}^T \mathbf{D} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{D}^{-1} \mathbf{g}$$
 (soluzione in Forma Chiusa)

Questa soluzione non è usata spesso nella tomografia ad emissione per le grandi dimensioni della matrice di sistema **H**, e spesso si preferisce **ricorrere a soluzioni iterative**.



Criterio ML

Introduzione

Scegliere come ricostruzione ottima, \hat{f} , quella che garantisce la massima probabilità di aver prodotto i dati misurati (massimizza la funzione di verosimiglianza p(g|f) o L(f)).

$$\hat{\mathbf{f}} = \arg\max_{f} p(\mathbf{g}|\mathbf{f})$$

Il criterio ML cerca la soluzione caratterizzata dalla maggiore vicinanza statistica alle osservazione misurate.



Criterio di massima verosimiglianza

Proprietà dello stimatore ML

- è asintoticamente non polarizzato (asymptotically unbiased), ossia al crescere del numero di osservazioni, la stima tende al valore vero $\left(E\left[\hat{f}\right] \to f\right)$
- è asintoticamente efficiente, ossia e lo stimatore a varianza minima, tra tutti gli stimatori nonpolarizzati. Detto altrimenti, lo stimatore ML e meno suscettibile al rumore presente nei dati.



Limite intrinseco degli approcci ML, LS e WLS

Sfortunamente per noi, nonostante la varianza (rumore) delle immagini ET ricostruite con stimatore ML sia la più bassa ottenibile con uno stimatore non polarizzato, questa continua ad essere ancora inaccettabilmente alta. Questo segue direttamente dalla loro tendenza a garantire la massima consistenza con i dati misurati: dato che i dati sono intrinsecamente rumorosi, l'immagine che è massimamente consistente con essi sarà per forza di cose rumorosa a sua volta.



Questo vale anche per stimatore LS e WLS!

(caso LS)
$$\begin{array}{l} \hat{\mathbf{f}} = \mathbf{H}^+\mathbf{g} = \mathbf{H}^+(\mathbf{H}\mathbf{f} + \mathbf{n}) & \vdots \\ = \mathbf{f} + \mathbf{H}^+\mathbf{n} & \vdots \end{array}$$

- Ipotesi di rumore additivo
- H nella tomografia ad emissione agisce come sistema passa-basso
- La sua inversa H^+ di conseguenza è passa alto



Possibili soluzioni

La soluzione consiste nell'accettare un certo *bias* nelle immagini ricostruite in cambio di una riduzione di varianza. Per far questo si introduce uno *smoothing* spaziale nelle immagini, che ovviamente va a ridurre il rumore a discapito della vicinanza della stima al valore vero.

Lo smoothing può essere introdotto in due modi:

- esplicitamente: applicando un filtraggio passa basso all'immagine ricostruita
- **implicitamente**: interrompendo prematuramente (prima di convergere all'effettiva soluzione ML) le iterazioni dell'algoritmo

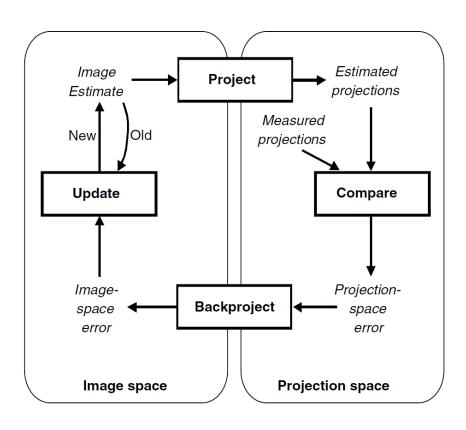


Approcci Bayesiani di stima regolarizzata (Maximum A Posteriori)

$$\hat{\mathbf{f}} = \arg \max_{f} [\ln p(\mathbf{g}|\mathbf{f}) + \ln p(\mathbf{f})]$$



Struttura generale di un algoritmo iterativo



stima corrente all'iterazione *n* $\hat{f}^{(n)}$

 $\widehat{f}^{(n)} \rightarrow \widehat{g}^{(n)}$

la stima corrente viene proiettata nel dominio dei sinogrammi

 $\hat{q}^{(n)} \Leftrightarrow q$

proiezione predetta viene confrontata con i valori misurati g

vettore di errore definito nello spazio proiettato

vettore di errore nello spazio immagine attraverso retroproiezione

 $\hat{f}^{(n+1)}$

stima aggiornata sulla base del feedback di errore



Derivazione ML-EM: missing data

L'approccio EM è standard per la soluzione di problemi ML in cui non si hanno a disposizione *tutti* i dati che renderebbero semplice la soluzione (**problema dei dati mancanti**)



Astrazione

ipotizziamo l'esistenza di un **set completo di dati S** in cui conosciamo sia il numero di fotoni misurati nel bin *j*, sia quanti, di questi, provengono dal pixel *k*

E-step

$$Q(f|\hat{f}^{(n)}) = E[\ln p(s|f)|g,\hat{f}^{(n)}]$$

→ Aspettazione della log-likelihood del set completo

M-step

$$\hat{\mathbf{f}}^{(n+1)} = argmax_f Q(\mathbf{f}|\hat{\mathbf{f}}^{(n)})$$

→ Massimizzazione della stessa



19

Derivazione ML-EM: E-step

 s_{im} = elemento del set completo di dati (variabile aleatoria Poissoniana)

$$g_{i} = \sum_{m} s_{im}$$

$$p(\mathbf{s}|\mathbf{f}) = \prod_{i} \frac{E[s_{im}]^{s_{im}} e^{-E[s_{im}]}}{s_{im}!}$$

$$E[s_{im}] = h_{im} f_{m}$$

$$\ln p(\boldsymbol{s}|\boldsymbol{f}) = \sum_{i} \sum_{m} [s_{im} \ln(h_{im} f_m) - h_{im} f_m - \ln(s_{im}!)]$$

$$Q(f|\hat{f}^{(n)}) = E[\ln p(s|f)|g, \hat{f}^{(n)}]$$
 E-step
= $\sum_{i} \sum_{m} \{E[s_{im}|g, \hat{f}^{(n)}] \ln(h_{im}f_{m}) - h_{im}f_{m} - E[\ln(s_{im}!)]\}$



E-step

Introduzione

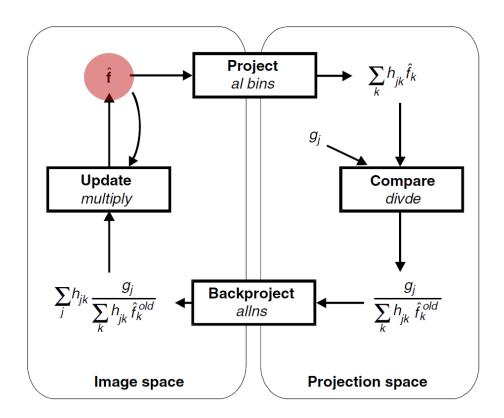
$$Q(\mathbf{f}|\hat{\mathbf{f}}^{(n)}) = \sum_{i} \sum_{m} E[s_{im} | \mathbf{g}, \hat{\mathbf{f}}^{(n)}] \ln(h_{im}f_{m}) - h_{im}f_{m} - E[\ln(s_{im}!)]$$

$$E[s_{im} | \mathbf{g}, \hat{\mathbf{f}}^{(n)}] = g_{i} \frac{h_{im}\hat{f}_{m}}{\sum_{m} h_{im}\hat{f}_{m}} \stackrel{\text{def}}{=} p_{im}$$

$$\frac{\partial Q(\boldsymbol{f}|\hat{\boldsymbol{f}}^{(n)})}{\partial f_j} = 0 = \sum_{i} \frac{p_{ij}}{\hat{f_i}^{(n+1)}} - h_{ij} \xrightarrow{\text{yields}}$$

 $\frac{\partial Q(\boldsymbol{f}|\hat{\boldsymbol{f}}^{(n)})}{\partial f_{j}} = 0 = \sum_{i} \frac{p_{ij}}{\hat{f_{j}}^{(n+1)}} - h_{ij} \quad \overset{\textbf{yields}}{\longrightarrow} \quad \hat{f_{j}}^{(n+1)} = \frac{\hat{f_{j}}^{(n)}}{\sum_{i'} h_{i'j}} \sum_{i} h_{ij} \frac{g_{i}}{\sum_{k} h_{ik} \hat{f_{k}}^{(n)}}$ M-step



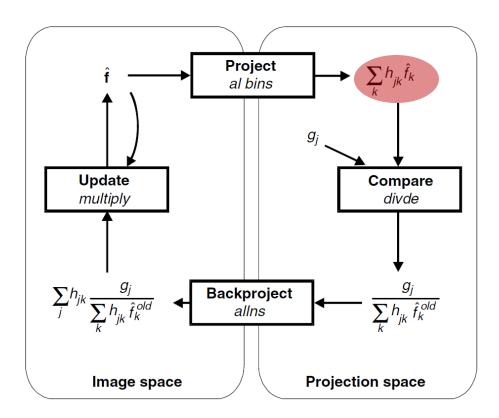


1) Stima iniziale uniforme dell'immagine

$$\hat{f_j}^{(0)}$$

$$\hat{f_j}^{(n+1)} = \frac{\hat{f_j}^{(n)}}{\sum_{i'} h_{i'j}} \sum_{i} h_{ij} \frac{g_i}{\sum_{k} h_{ik} \, \hat{f_k}^{(n)}}$$



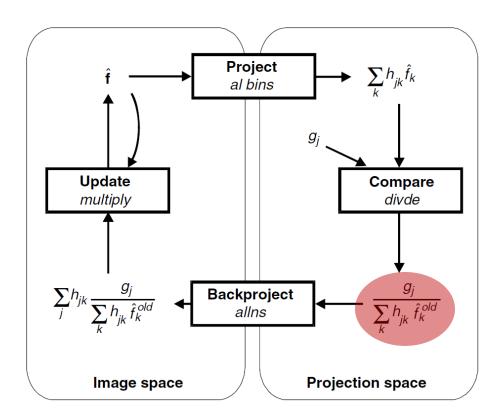


2) Simulazione della misura a partire dalla stima (proiezione)

$$\hat{g_i}^{pred} = \sum_{k} h_{ik} \, \hat{f_k}^{(n)}$$

$$\hat{f_j}^{(n+1)} = \frac{\hat{f_j}^{(n)}}{\sum_{i'} h_{i'j}} \sum_{i} h_{ij} \frac{g_i}{\sum_{k} h_{ik} \hat{f_k}^{(n)}}$$



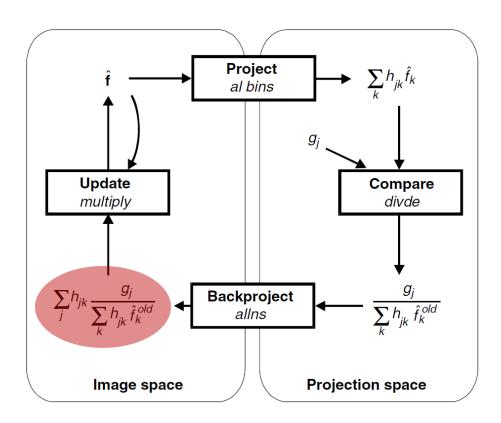


3) Confronto con la misura vera

$$e_i^{(g)} = \frac{g_i}{\hat{g}_i^{pred}}$$

$$\hat{f_j}^{(n+1)} = \frac{\hat{f_j}^{(n)}}{\sum_{i'} h_{i'j}} \sum_{i} h_{ij} \frac{g_i}{\sum_{k} h_{ik} \hat{f_k}^{(n)}}$$





Passaggio al dominio immagine (backprojection)

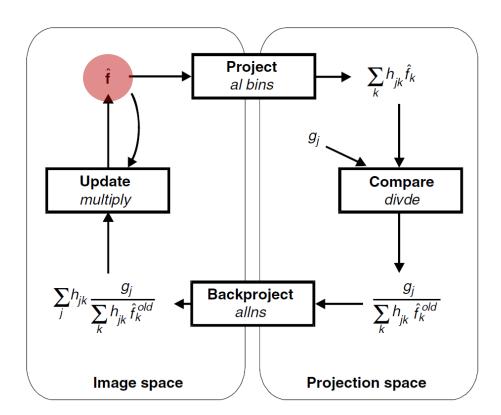
$$e_i^{(f)} = \sum_i h_{ij} \, e_i^{(g)}$$

$$\hat{f}_{i}^{(n+1)} = \frac{\hat{f}_{j}^{(n)}}{\sum_{h \in \mathcal{I}} g_{i}}$$



Introduzione

ML-EM



5) Aggiornamento della stima (pesatura della stima precedente)

$$\hat{f_j}^{(n+1)} = \frac{\hat{f_j}^{(n)}}{\sum_i h_{ij}} e_i^{(f)}$$

ML-EM

$$\hat{f_j}^{(n+1)} = \frac{\hat{f_j}^{(n)}}{\sum_{i'} h_{i'j}} \sum_{i} h_{ij} \frac{g_i}{\sum_{k} h_{ik} \hat{f_k}^{(n)}}$$



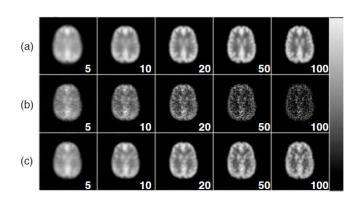
Algoritmo ML-EM: numero di iterazioni

1) Convergenza garantita e predicibile ma lenta

- Dati rumorosi: almeno 30, 50 iterazioni necessarie
- Tempi di esecuzione circa 2 ordini di grandezza superiori alla FBP

2) Immagini finali rumorose

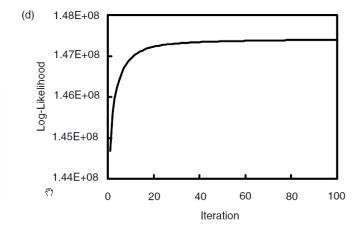
Tende alla massima consistenza con misure rumorose ...



Nota bene ...

Introduzione

La ML-EM ricostruisce prima le **frequenze spaziali più basse**, per poi sviluppare gradualmente le componenti ad alta frequenza: bloccare le iterazioni equivale quindi implicitamente ad un **filtraggio passa-basso** dell'immagine ricostruita.

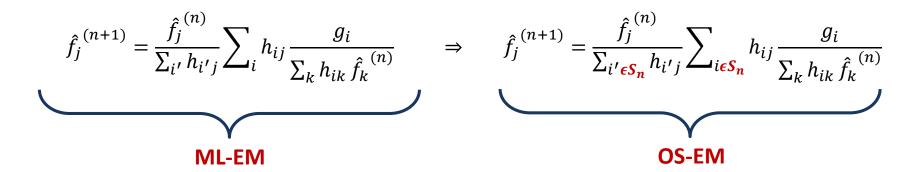




Esercitazione

Introduzione

Algoritmo OS-EM: accelerare la convergenza



La retroproiezione viene applicata soltanto ai bin che appartengono al sotto insieme (subset) S_n del sinogramma. Ad ogni aggiornamento, un diverso subset viene preso in considerazione in quella che viene definita sotto-iterazione dell'algoritmo (mentre un intero passaggio di tutti i subset costituisce un'iterazione vera e propria):

- $N_{subset} = 1$ \rightarrow ci riconduciamo al caso ML-EM
- $N_{subset} = N_{proiezioni}$ \rightarrow otteniamo un algoritmo simile alla (M)ART

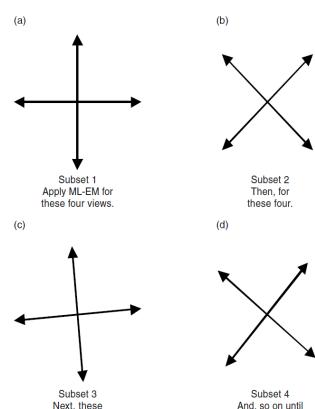


Algoritmo OS-EM: organizzazione dei subsets

$$\hat{f}_{j}^{(n+1)} = \frac{\hat{f}_{j}^{(n)}}{\sum_{i' \in S_n} h_{i'j}} \sum_{i \in S_n} h_{ij} \frac{g_i}{\sum_k h_{ik} \hat{f}_k^{(n)}}$$

PRINCIPALI REGOLE DI SCELTA

- In ogni subset deve comparire il contributo di tutti i pixel del FOV
- Le proiezioni che appartengono ad un subset sono scelte in modo da avere massima distanza angolare possibile tra loro
- Anche l'alternanza nel processing dei subset cerca di massimizzare la differenza tra subset successivi





$$f \xrightarrow{3 \times 3} \begin{bmatrix} f_1 & f_4 & f_7 \\ f_2 & f_5 & f_8 \\ f_3 & f_6 & f_9 \end{bmatrix}$$

Introduzione

$$H_{24x9}$$

 θ_3

 θ_4

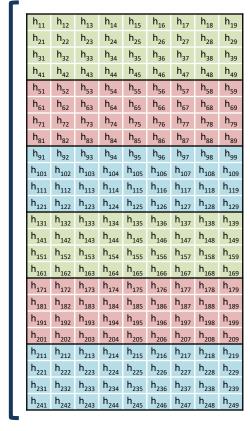
N_{proiezioni}: 6

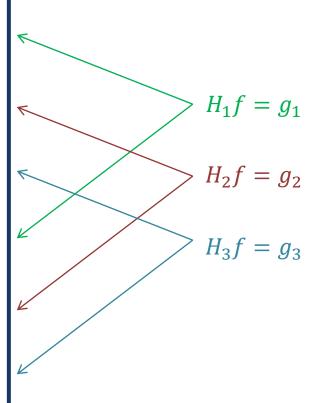
N_{posizioni_angolari}: 4

$$g \rightarrow [6x4]$$

 θ_5









OS-EM vs ML-EM

- ➤ Il numero di subset scelto determina il **grado di accelerazione** della OS-EM
- ightharpoonup (Empiricamente) dopo N iterazioni di OS-EM si raggiunge un **punto di convergenza** paragonabile a quello ottenuto con $(N_{subset} \times N)$ iterazioni di ML-EM
- ➤ Il rumore cresce a velocità comparabile e l'algoritmo va arrestato (ed eventualmente filtrato) ancora più precocemente
- Nonostante la somiglianza con la ML-EM, la OS-EM non è realmente un algoritmo EM e quindi non c'è prova teorica della sua convergenza



Correzioni iterative (ML-EM e OS-EM)

Abbiamo detto precedentemente che se il sistema di imaging corregge internamente i sinogrammi prima che avvenga la ricostruzione, la statistica muta e non sarebbe più valido lavorare nell'ipotesi di Poissonianità dei dati su cui si basno ML-EM ed OS-EM.

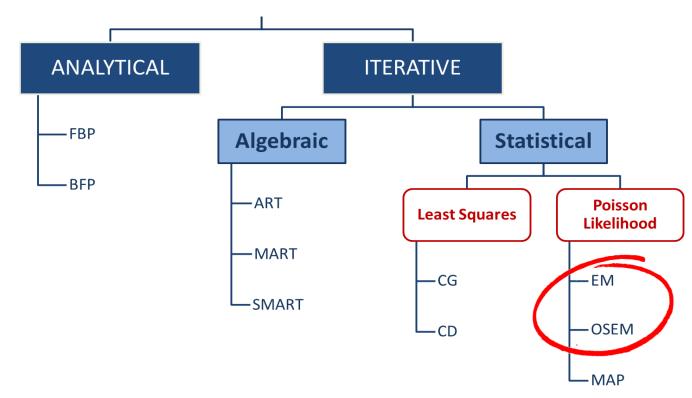
Tra i vantaggi del meccanismo di *feedback* che è alla base di queste tecniche iterative c'è però anche la possibilità di **incorporare delle stime delle sorgenti di disturbo** (le stesse che in maniera inversa verrebbero usate per pre-correggere i sinogrammi) **all'interno dello step di proiezione**, riuscendo quindi a ricostruire e corregge le immagini allo stesso tempo.

$$\hat{f}^{(n+1)} = \frac{\hat{f}^{(n)}}{H^T 1} H^T \frac{g}{att.* \left(H\hat{f}^{(n)}\right) + scatt + rand}$$
 ML-EM

OS-EM
$$\begin{cases} \hat{f_k}^{(n+1)} = \frac{\hat{f_k}^{(n)}}{{H_k}^T 1_k} {H_k}^T \frac{g_k}{att_k.* \left({H_k} \hat{f_k}^{(n)}\right) + scatt_k + rand_k} \end{cases}$$



Reconstruction Methods

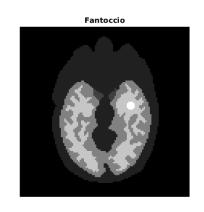


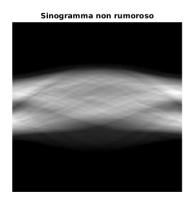


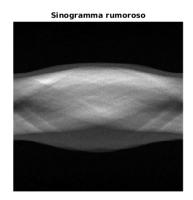
Esercitazione (0)

MATERIALE

- Fantoccio cerebrale 2D (111x111 pixel) preparato (file brain.mat)
- Funzione MATLAB *Calcolo_A.m* già utilizzata nell'esercitazione sulle tecniche di ricostruzione analitiche
- Scheletro dell'esercitazione da svolgere
- Scheletro della funzione Calcolo_Ablock.m con descrizione di quale debba essere la sua struttura e il suo funzionamento.





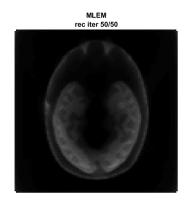


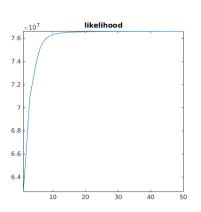


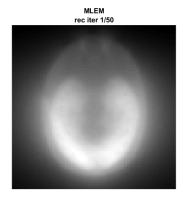
Esercitazione (1)

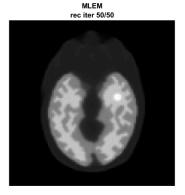
1. Ricostruzione ML-EM

- a) Implementare l'algoritmo ML-EM per la ricostruzione del sinogramma rumoroso generato ai punti precedenti.
- b) Valutare la qualità della ricostruzione a seconda che vengano corretti o meno i disturbi simulati.
- Salvare l'immagine intermedia ricostruita ad ogni iterazione in un vettore 3D (N,N,iter mlem)







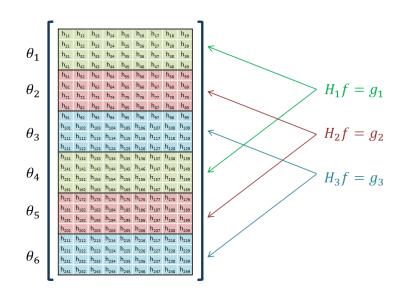




Esercitazione (2)

2. Ricostruzione OSEM: calcolo dei blocchi della matrice di sistema

- a) Creare una funzione esterna (partire dal le Calcolo Ablock.m fornito) per l'estrazione dei blocchi della matrice di sistema con cui ricostruire i singoli subset del sinogramma.
- b) La descrizione della funzione e fornita nel le dedicato: e importante assicurarsi che restituisca in output 'nblock' segmenti della matrice di sistema A e, per ciascuno di essi, tenga traccia delle proiezioni che fanno parte del subset a cui e associato un determinato blocco.

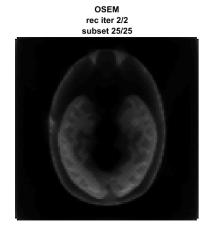


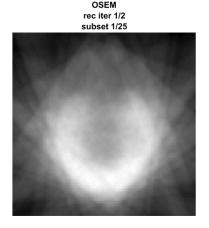


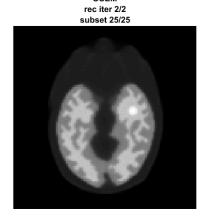
Esercitazione (3)

1. Ricostruzione OS-EM

- a) Implementare l'algoritmo OS-EM per la ricostruzione del sinogramma rumoroso generato ai punti precedenti.
- b) Valutare la qualità della ricostruzione a seconda che vengano corretti o meno i disturbi simulati.
- c) Salvare l'immagine intermedia ricostruita ad ogni iterazione in un vettore 3D (N,N,iter_mlem*nblock)









iter: 1| subs: 16

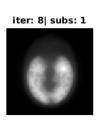
Introduzione

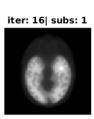
4. Analisi della relazione tra numero di subset e numero di iterazioni OSEM

Visualizzare il risultato della ricostruzione OSEM usando diverse combinazioni di valori per il numero di subset (fattore accelerazione) e il numero di iterazioni di ricostruzione.

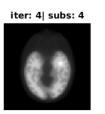


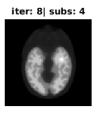


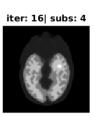


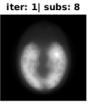


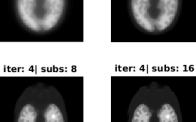


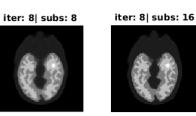


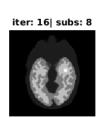
















Esercitazione (4)

5. Valutazione dell'accelarazione ottenuta grazie all'algoritmo OS-EM

Misurare i tempi di esecuzione di una ricostruzione ML-EM (totale, non delle singole sub-iterazioni) e di una ricostruzione OS-EM e vericare che ci sia una velocizzazione del processo di ricostruzione a parita di iterazioni utilizzate.

NOTA BENE:

- mettersi nella condizione iter mlem = iter osem*nblock
- vericare tutte e 3 le combinazioni possibili
- assicurarsi di misurare l'effettivo tempo di ricostruzione

ART 180 subset 180 iter	5.334660e+00
MLEM 360 iter	3.609439e+00
OSEM 9 subset 40 iter	4.635321e-01
OSEM 40 subset 9 iter	1.287901e-01



Esercitazione (4)

FACOLTATIVO

Introduzione

1. Ricostruzione ART

- a) Utilizzare la funzione creata Calcolo_Ablock.m per precalcolare la struttura in blocchi della matrice utilizzata per la ricostruzione ART.
- b) Seguire il modello dell'implementazione fornita per l'algoritmo OSEM e produrre una implementazione ottimizzata dell'algoritmo ART sviluppato nella precedente esercitazione
- c) Valutare il tempo necessario alla ricostruzione con questo particolare algoritmo
- 2. Ripetere il punto (5) valutando l'andamento dei tempi di calcolo in funzione del numero di iterazioni richieste ai vari algoritmi.

