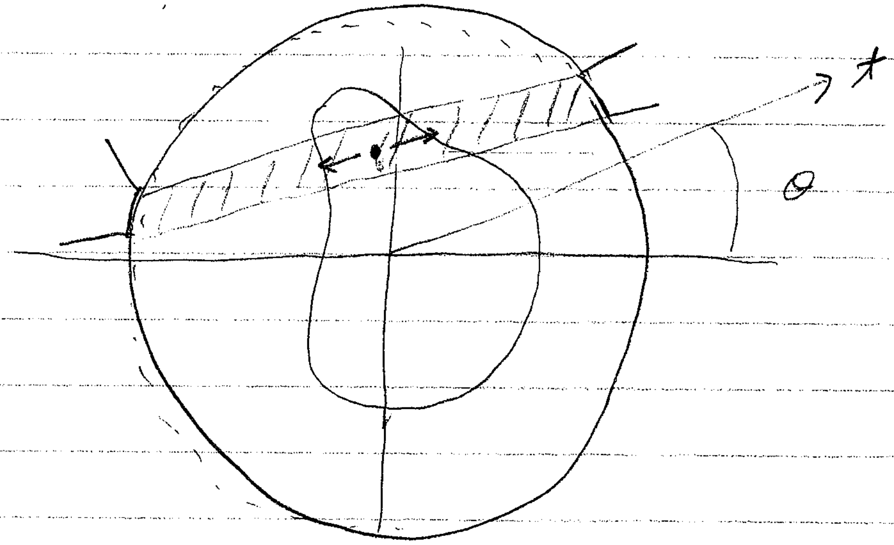
EMISSION TOMOGRAPHY



1. Inject radioactive dye
2. Simultaneously measure projections at all angles θ and displacements t

Observed data

$y_i = y(\theta_i, t_i)$
$\quad$ = sum of emissions at angle θ_i displacement t_i

Unknown image

x_j = <u>Rate</u> of emission from j^{th} pixel

Lecture 33

- Normal assumptions:
  1) y_i are conditionally independent given X
  2) y_i is Poisson distributed with

$$E[Y] = \underset{\uparrow\ \text{projection matrix}}{A} X$$

$$p(y|x) = \prod_{i=1}^{M} p(y_i|x)$$

$$= \prod_{i=1}^{M} \frac{(A_{i*}X)^{y_i} e^{-A_{i*}X}}{y_i!}$$

$$\underset{\uparrow\ i^{th}\text{ row}}{A_{i*}} X \triangleq \sum_j A_{ij} x_j$$

$$\log p(y|x) = \sum_{i=1}^{M} \left(y_i \log(A_{i*}X) - A_{i*}X - \log(y_i!) \right)$$

- Concave function of X

$$\hat{X}_{MAP} = \arg\min_{X} \left\{ \sum_{i=1}^{M} \left(A_{i*}X - y_i \log(A_{i*}X) \right) + \frac{1}{2} \sum_{\{i,j\} \in C} b_{i-j}\, \rho\left(|x_i - x_j| / \sigma \right) \right\}$$

How do we solve this?

1) Direct optimization (convex)

2) EM algorithm ⇐ traditional approach

EM algorithm

define the missing data

Z_{ij} - number of photons detected by y_i and admitted by x_j

- Can be shown that Z_{ij} are conditionally independent given X and Poisson

$$p(Z_{ij} | x_j) = \frac{(A_{ij} x_j)^{Z_{ij}} e^{-A_{ij} x_j}}{Z_{ij}!}$$

$$y_i = \sum_j Z_{ij}$$

- Bayesian approach to EM
  - X is like θ
  - Z includes Y since $Y_i = \sum_j Z_{ij}$

$$Q(\hat{x}, x) = E[\log p(Z | \hat{x}) | Y = y, x] + \log p(\hat{x})$$

$$= E[\log p(Z, \hat{x}) | Y = y, x]$$

Lecture 34

$$\log p(z/\hat{x}) = \sum_{i,j} \left\{ z_{ij} \log(A_{ij}\hat{X}_j) - A_{ij}\hat{X}_j - \log(z_{ij}!) \right\}$$

$$Q(\hat{x}, x) = \sum_{ij} \left\{ \bar{Z}_{ij} \log(A_{ij}\hat{X}_j) - A_{ij}\hat{X}_j \right\} + C + \log p(\hat{x})$$

$$\bar{Z}_{ij} = E[z_{ij} / Y = y, X]$$

$$= E[z_{ij} / Y_i = y_i, X]$$

$$= E[z_{ij} / \sum_p z_{ip}, X]$$

$$p(z_{ij} / \sum_p z_{ip}, X) = C\, p(\sum_{p \neq i} z_{ip} / z_{ij}, x) \cdot p(z_{ij}/x)$$

$$\lambda_2 \triangleq \sum_{p \neq i} a_{ip} \lambda_p \qquad \lambda_1 = a_{ij} x_j$$

$$= C \frac{\lambda_2^{y_i - z_{ij}} e^{-\lambda_2}}{(y_i - z_{ij})!} \frac{\lambda_1^{z_{ij}} e^{-\lambda_1}}{z_{ij}!}$$

$$= \binom{Y_i}{z_{ij}} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^{z_{ij}} \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{y_i - z_{ij}}$$

Binomial

$$\bar{Z}_{ij} = \frac{\lambda_1}{\lambda_1 + \lambda_2} y_i = \frac{a_{ij} x_j}{\sum_p a_{ip} x_p} y_i$$

$Q(\hat{x}, x) =$

$$= \sum_{i,j} \left\{ \frac{y_i a_{ij} x_j}{\sum_p a_{ip} x_p} \ln(a_{ij} \hat{x}_j) - a_{ij} \hat{x}_j \right\}$$

$$- \frac{1}{2} \hat{x}^t B \hat{x}$$

$$\hat{\hat{x}} = \arg\max_{\hat{x}} Q(x, \hat{x})$$

How do we maximize this?

- <u>2 approaches</u>

- Levitan and Herman
  - $B = \frac{1}{\sigma^2} I \Rightarrow$ independent pixels
  - no spatial correlation
  - not very good

- Hebert and Leahy
  - Generalized EM algorithm (GEM)

$$Q(x^{(k+1)}, x^{(k)}) > Q(x^{(k)}, x^{(k)})$$

  - Use 1 pass of coordinate decent to maximize Q
  - Nested iterations

Direct Optimization Approach

- Direct Coordinatewise optimization of $\log p(x|y)$
- Quadratic approximation
- Convergence analysis
- Exact optimization

$$L(y|x) \triangleq \log p(y|x)$$

$$= \sum_{i=1}^{M} \left(y_i \log(A_{i*}x) - A_{i*}x - \log y_i! \right)$$

$$= -\sum_{i=1}^{M} \underbrace{f_i}_{\text{convex}}(A_{i*}x)$$

Define $\tilde{p}_i = A_{i*}x$

$$L(y|x) = -\sum_{i=1}^{M} f_i(\tilde{p}_i)$$

$$L(y|x) \cong c + b(\tilde{p}-y_i) - \frac{1}{2}(\tilde{p}-y)^t D(\tilde{p}-y)$$

- It can be shown that

$$b_j = \frac{\partial}{\partial \tilde{p}_j}\left(-\sum_{i=1}^{M} f_i(\tilde{p}_i)\right)\Bigg|_{\tilde{p}_i=y_i} = 0$$

$$D_{jk} = \frac{\partial^2}{\partial \tilde{p}_j \partial \tilde{p}_k}\left(\sum_{i=1}^{M} f_i(\tilde{p}_i)\right)\Bigg|_{\tilde{p}_i=y_i} = \frac{\delta_{jk}}{y_j}$$

$$D = \text{diag}\left(1/y_1, 1/y_2, \ldots, 1/y_M \right)$$

$$\log p(y|x) \cong -\frac{1}{2}(\tilde{p}-y)^t D(\tilde{p}-y) + c$$

$$= -\frac{1}{2}(y-Ax)^t D(y-Ax) + c$$

$$\hat{x}_{MAP} = \arg\min_x \left\{ \frac{1}{2}(y-Ax)^t D(y-Ax) + \frac{1}{2}x^t B x \right\}$$

- Coordinate Decent algorithm

  - Must keep a state vector ⇐ <u>important</u>

$$e = y - Ax$$

change the i^{th} pixel $\left\{ x^{(k+1)} = \begin{cases} x_j^{(k)} & j \neq i \\ x_i^{(k)} + \Delta x & j = i \end{cases} \right.$

$$L(y|x^{(k+1)}) = L(y|x^{(k)}) + \theta_1 \Delta x + \frac{\theta_2}{2}(\Delta x)^2$$

$$\theta_1 = -A_{*i}^t De$$

$$\theta_2 = +A_{*i}^t D A_{*i} \Leftarrow \text{precompute for each pixel } i$$

$$e^{(k+1)} = e^{(k)} - A_{*i}\Delta x$$

$$x_i^{(k+1)} = \arg\min_{x_i} \left\{ \theta_1 (x_i - x_i^{(k)}) + \frac{\theta_2}{2} (x_i - x_i^{(k)})^2 + \frac{1}{2} \sum_{j \in \partial i} g_{i-j} (x_i - x_j^{(k)})^2 \right\}$$

where $B_{ij} = \delta_{i-j} - g_{i-j}$

$$\sum_j B_{ij} = 1$$

- Computation

  - Most of A is empty
    $A_{ji} \neq 0 \Leftrightarrow$ projection j intersects pixel i

  - Average pixel has M_o projections $>> 1$

  - θ_2 can be precomputed

  - θ_1 requires $2M_o$ multiplies

  Computation per iteration $\cong 2M_o N$

Example: 128 × 128 crossection image

$$\begin{cases} 128 \text{ projection angles} \\ 128 \text{ projection displacements} \end{cases}$$

Typical value for M_0 is

$$M_0 = 2\,(128) = 256$$

Lecture 35

Convergence Behavior

- Gradient Decent

$$\hat{x}_{MAP} = \arg\min_x \underbrace{\left\{ \tfrac{1}{2}(y-Ax)^t D(y-Ax) + \tfrac{1}{2} x^t B x \right\}}_{L(y,x)}$$

$$X^{(k+1)} = X^{(k)} + \mu \nabla_x L(y, x)$$

$$= x^{(k)} - \mu\left(A^t D (A X^{(k)} - y) + B X^{(k)}\right)$$

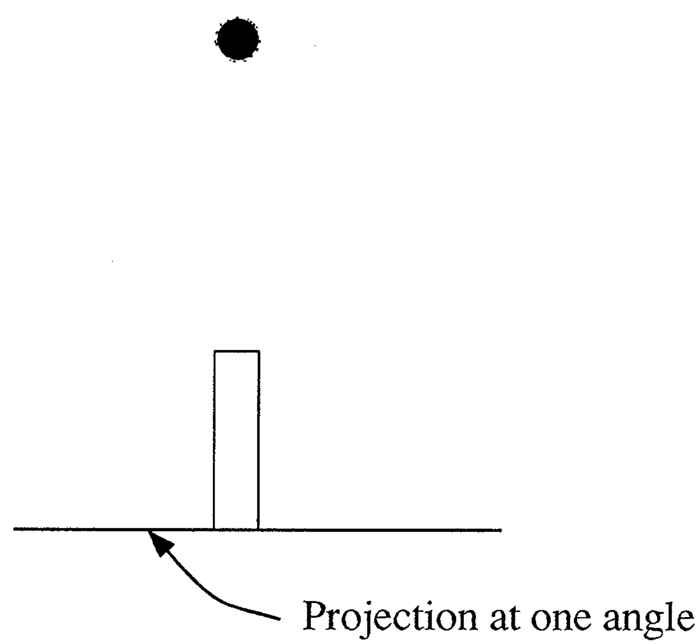
$$X^{(k+1)} = \left(I - \mu(A^t DA + B)\right) X^{(k)} + \underbrace{\alpha A^t D y}_{\text{forcing function}}$$

convergence is determined by the homogeneous difference equation

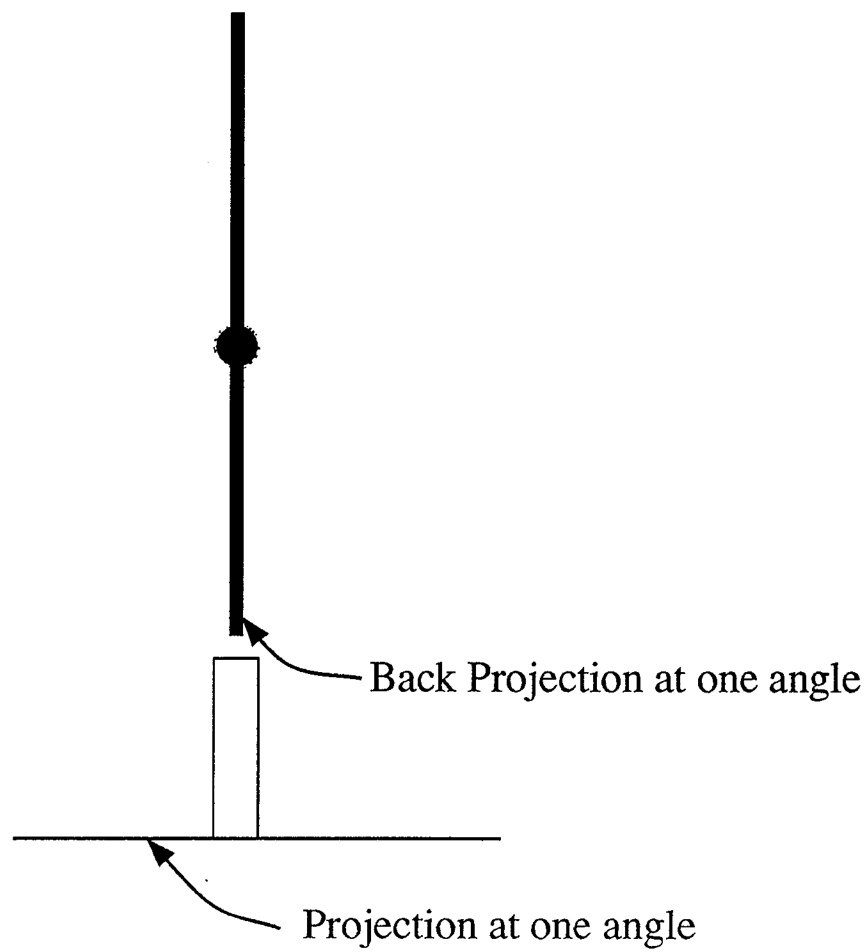
$$e^{(k+1)} = \left(I - \mu(A^t DA + B)\right) e^{(k)}$$

$$X^{(k)} = \hat{X}_{MAP} + e^{(k)}$$

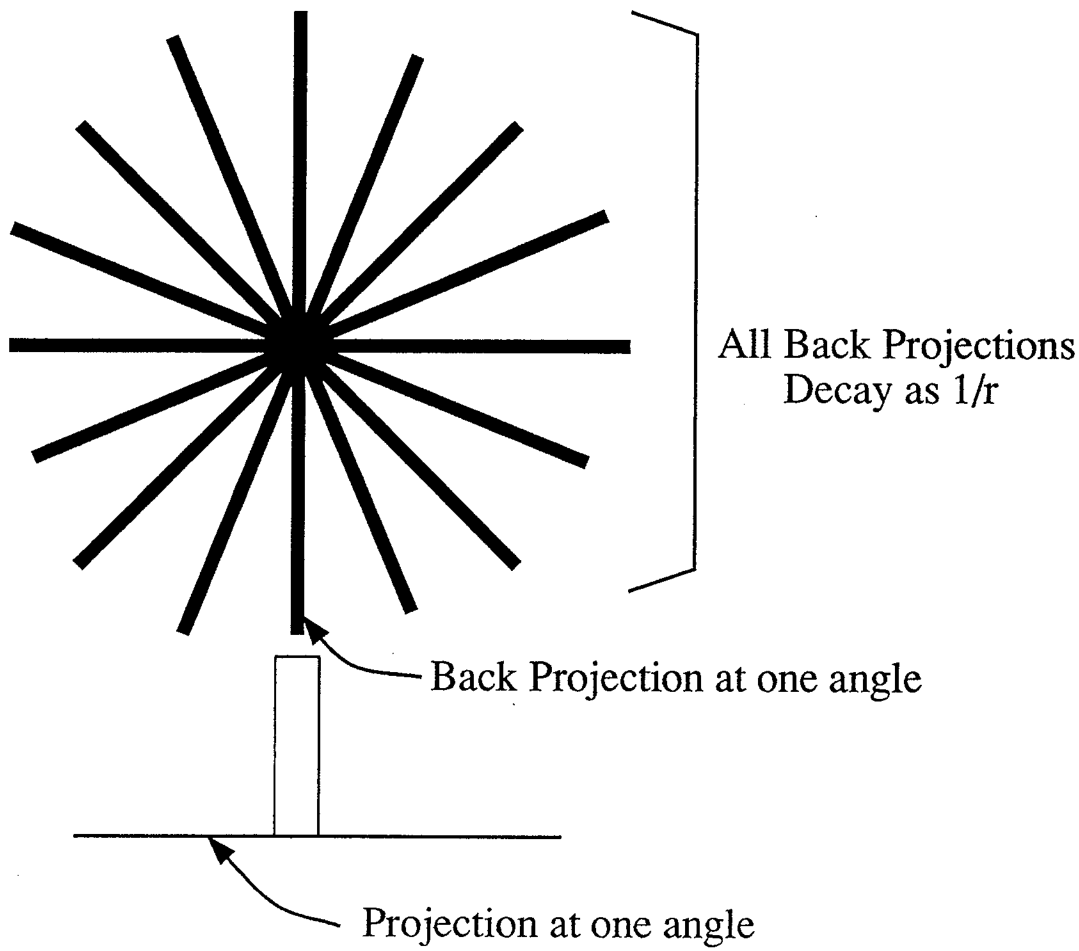
Projection/Back Projection of an Impulse



Projection/Back Projection of an Impulse



Projection/Back Projection of an Impulse



Structure of $A^t D A$

- assume $D = \frac{1}{\sigma_d^2} I$
- $(A^t A)_{sr}$ for $s = (s_1, s_2)$ $r = (r_1, r_2)$
- Can be shown that

$$(A^t A)_{sr} \cong \frac{\beta}{\sqrt{(s_1 - r_1)^2 + (s_2 - r_2)^2}}$$

- Muliplication by $A^t A$ is convolution

$$A^t A X \overset{\sim}{\Longleftrightarrow} H(\omega) X(\omega)$$

- $H(\omega) = \mathcal{F}\left\{ \frac{\beta}{r} \right\} = \frac{\beta}{\|\omega\|}$

$$e^{(k+1)}(\omega) = \left(1 - \mu\left(H(\omega) + \frac{1}{\sigma^2}(1 - G(\omega))\right)\right) e^{(k)}(\omega)$$

$$= P(\omega)\, e^{k}(\omega)$$

$$e^{(k)}(\omega) = (P(\omega))^{k} e^{(0)}(\omega)$$

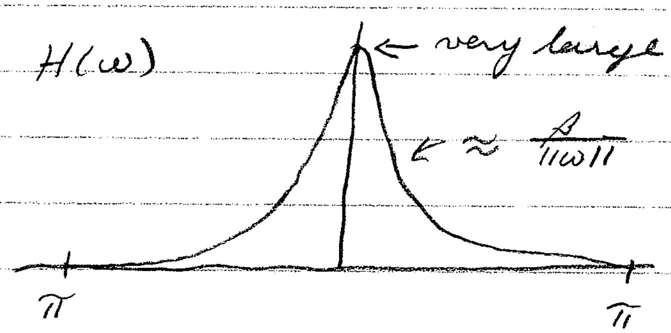
Lecture 36

- Stable convergence $\Leftrightarrow$ $|P(\omega)| < 1 \quad \forall \omega$

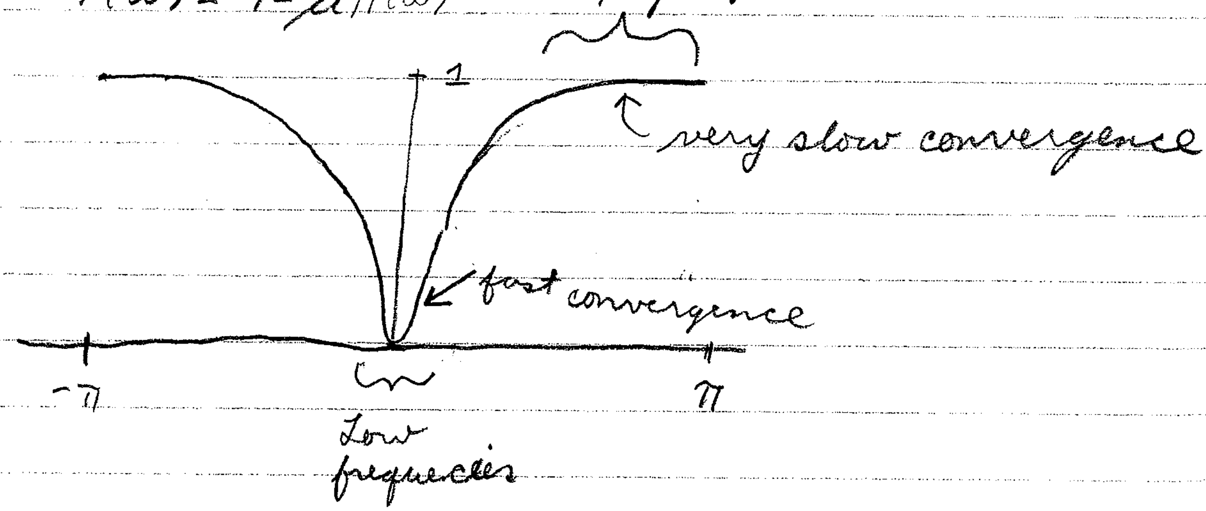
$$\left| 1 - \mu \Big(\underbrace{H(\omega)}_{\text{low pass}} + \frac{1}{\sigma^2} \underbrace{(1 - G(\omega))}_{\text{High pass}} \Big) \right| < 1$$

Consider $\sigma^2 \to \infty$ (no prior)

$$|1 - \mu H(\omega)| < 1 \Rightarrow 0 < \mu < \frac{2}{\max_{\omega} H(\omega)}$$



$P(\omega) = 1 - \mu H(\omega)$



Coordinate Decent

$$\frac{\partial}{\partial x_i}\left(\frac{1}{2}(y-Ax)^t D(y-Ax) + \frac{1}{2} x^t B x \right) = 0$$

$$= (A^t D A x - A^t D y + B x)_i$$

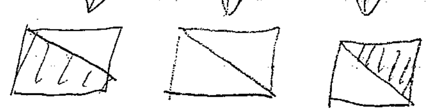
$$= (H_{i*} + B_{i*})\, x - b_i$$

where $b = A^t D y$

$H = A^t D A$

Define

$$\underset{\text{lower}}{L} + \underset{\text{diag}}{K} + \underset{\text{upper}}{U} = H + B$$



$$L x^{(k+1)} + K x^{(k+1)} + U x^{(k)} = H + B$$

- $x^{(k)}$ is the k^{th} <u>complete</u> replacement

$$\boxed{x^{(k+1)} = -(K+L)^{-1} U x^{(k)} + (K+L)^{-1} b}$$

Lecture 37

$$X^{(k+1)}(\omega) = \frac{-U(\omega)}{K(\omega)+L(\omega)} X(\omega) + \frac{b(\omega)}{K(\omega)+L(\omega)}$$

Note that since H and B are Toeplitz
$\Rightarrow K(\omega) = K_0$

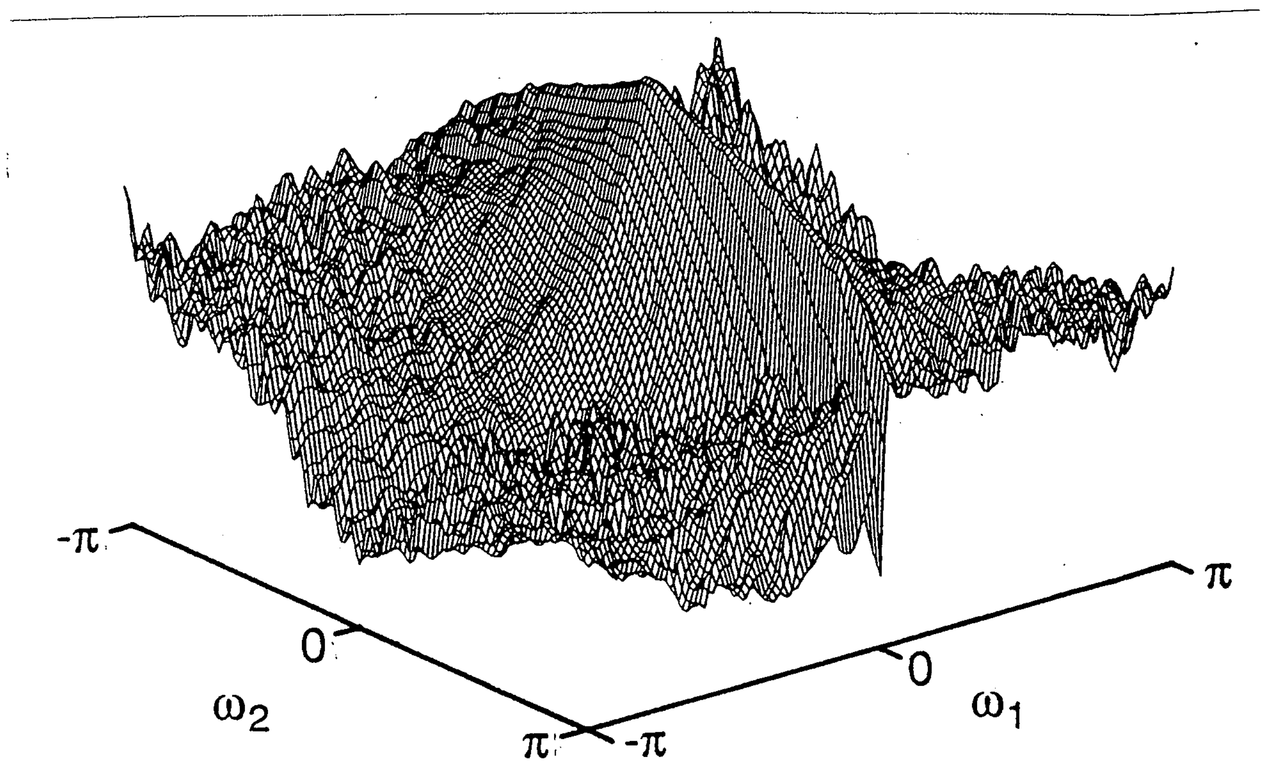
- Homogeneous solution

$$e^{(k)}(\omega) = \left(\frac{-U(\omega)}{K_0 + L(\omega)}\right)^k e^{(0)}(\omega)$$

$$P(\omega) = \frac{-U(\omega)}{K_0 + L(\omega)}$$

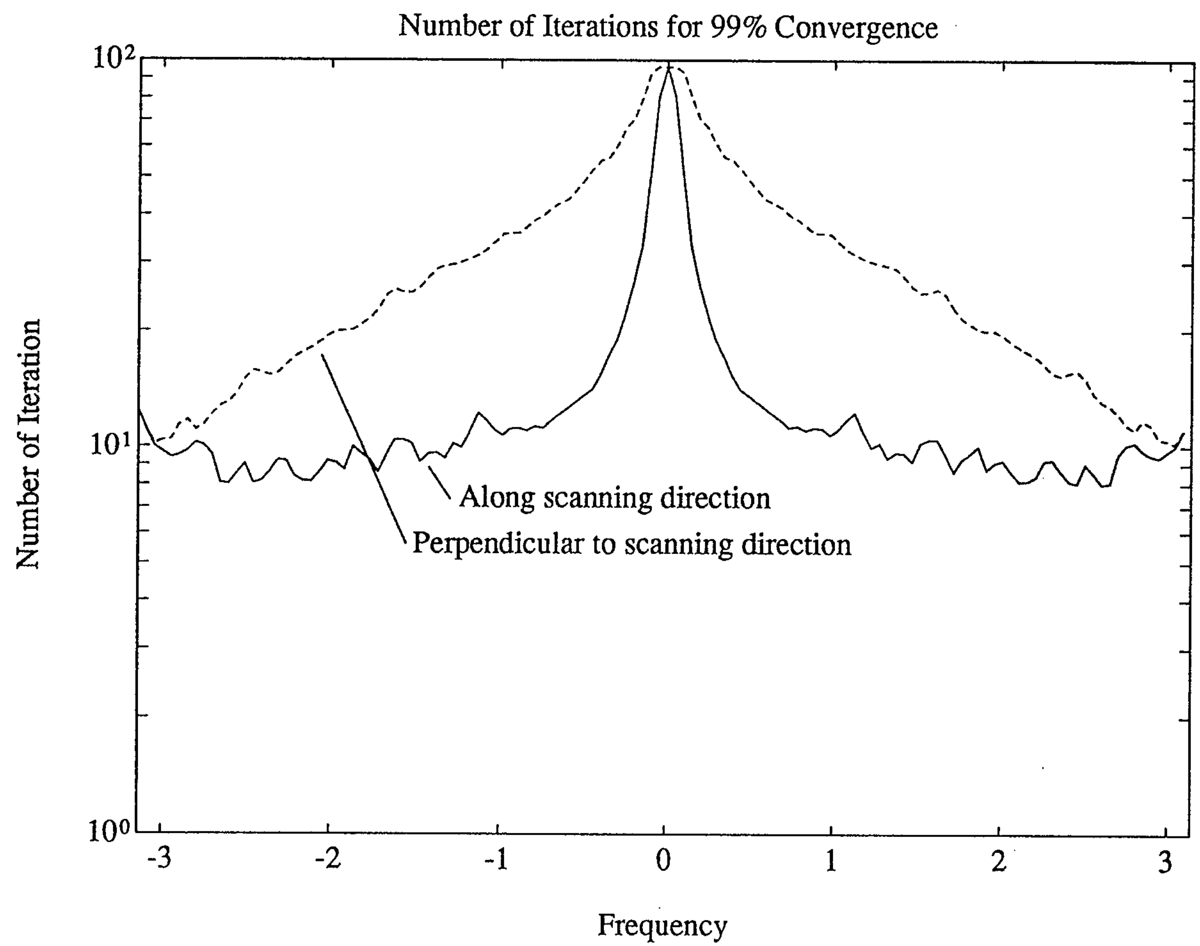
Convergence Rate

- The two dimensional convergence rate .



Convergence Rate

- Worst case convergence in each axis



Lecture 37

Exact MAP Optimization

- Coordinate decent

$$L(y|x) = \sum_{i=1}^{M} \left(y_i \log(A_{i*}X) - A_{i*}X - \log y_i! \right)$$

- We must update $X_j = \alpha$

$$g(\alpha) = \sum_{i=1}^{M} \left(y_i \log\left(A_{ij}(\alpha - X_j) + A_{i*}X \right) - A_{ij}\alpha \right) + C$$

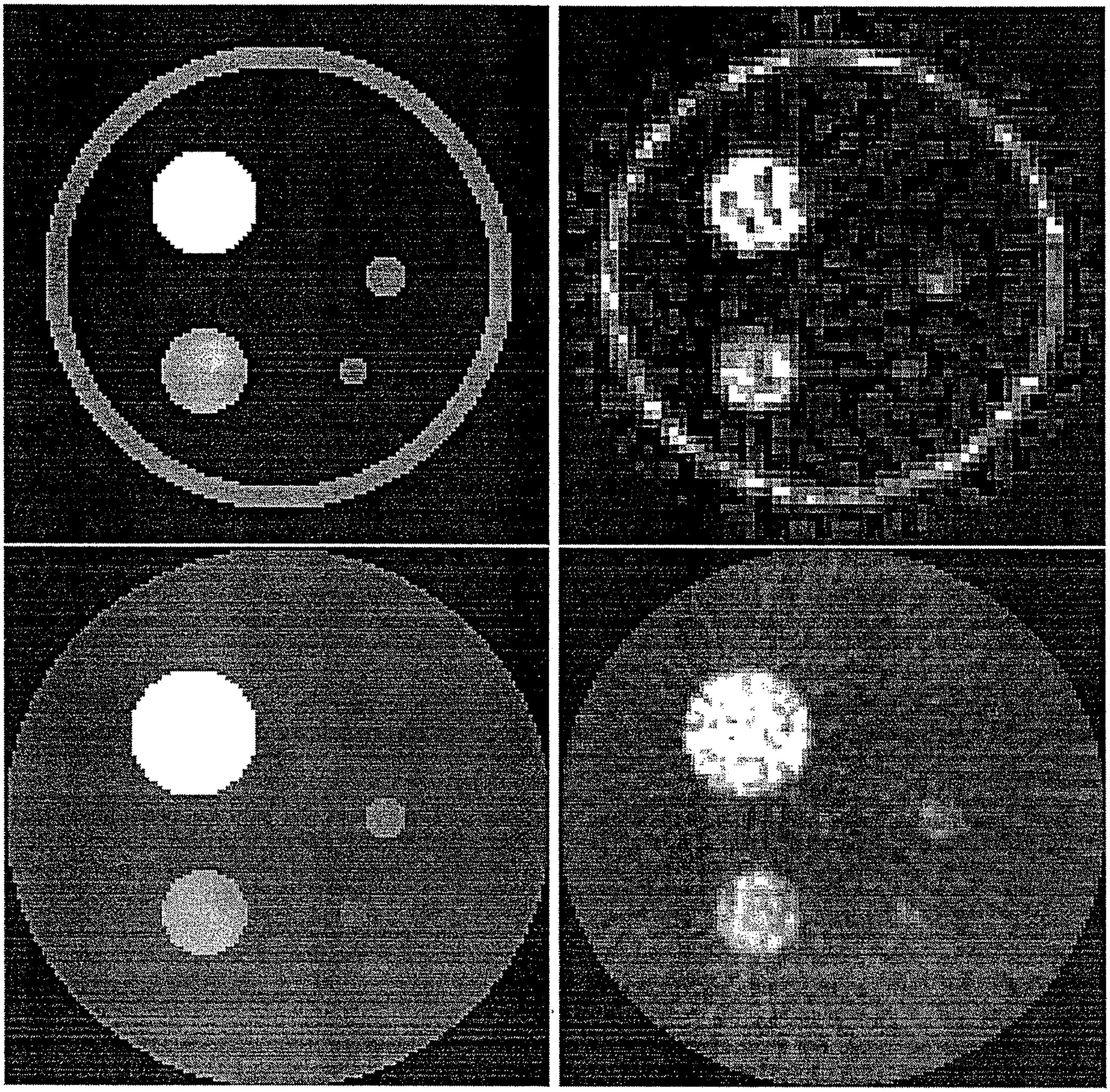
$$\theta_1 \triangleq \left. \frac{d\,g(\alpha)}{d\alpha} \right|_{\alpha = X_j}$$

$$= \sum_{i=1}^{M} P_{ij} \left(1 - \frac{y_i}{\tilde{p}_i} \right)$$

$$\theta_2 = \sum_{i=1}^{M} -y_i \left(\frac{P_{ij}}{\tilde{p}_i} \right)^2$$

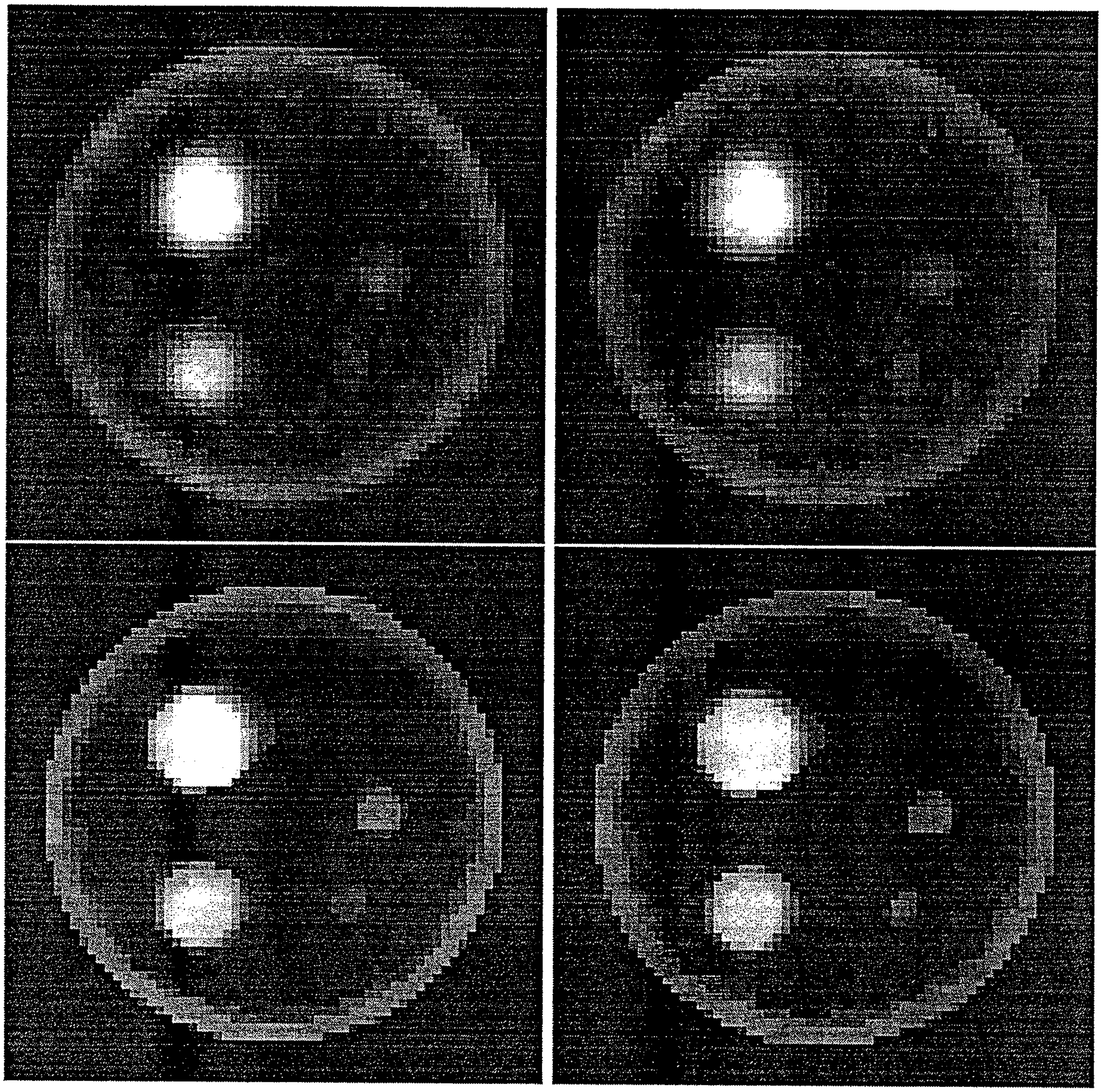
where $\tilde{p} = AX$

$$X_i^{(k+1)} = \arg\min_{\alpha} \left\{ \theta_1 (\alpha - X_i^{(k)}) + \frac{\theta_2}{2} (\alpha - X_i^{(k)})^2 + \sum_{j \in \partial i} b_{ij} (\alpha - X_j^{(k)})^2 \right\}$$



a	b
c	d

Figure 2: Original synthetic phantoms and their CBP reconstructions for emission and transmission examples. (a) Emission phantom with higher emission intensities in lighter areas, and (b) CBP emission reconstruction. (c) Transmission phantom with higher density in lighter areas, and (d) CBP transmission reconstruction. CBP reconstructions were computed using a raised cosine rolloff filter, and served as the initial estimate for the iterative statistical methods.



a	b
c	d

Figure 4: Emission MAP reconstructions with a Gaussian MRF prior, and (a) exact log likelihood function; (b) quadratic approximation. MAP estimates resulting from GGMRF model with $q = 1.1$, and (c) exact log likelihood; (d) quadratic approximation.

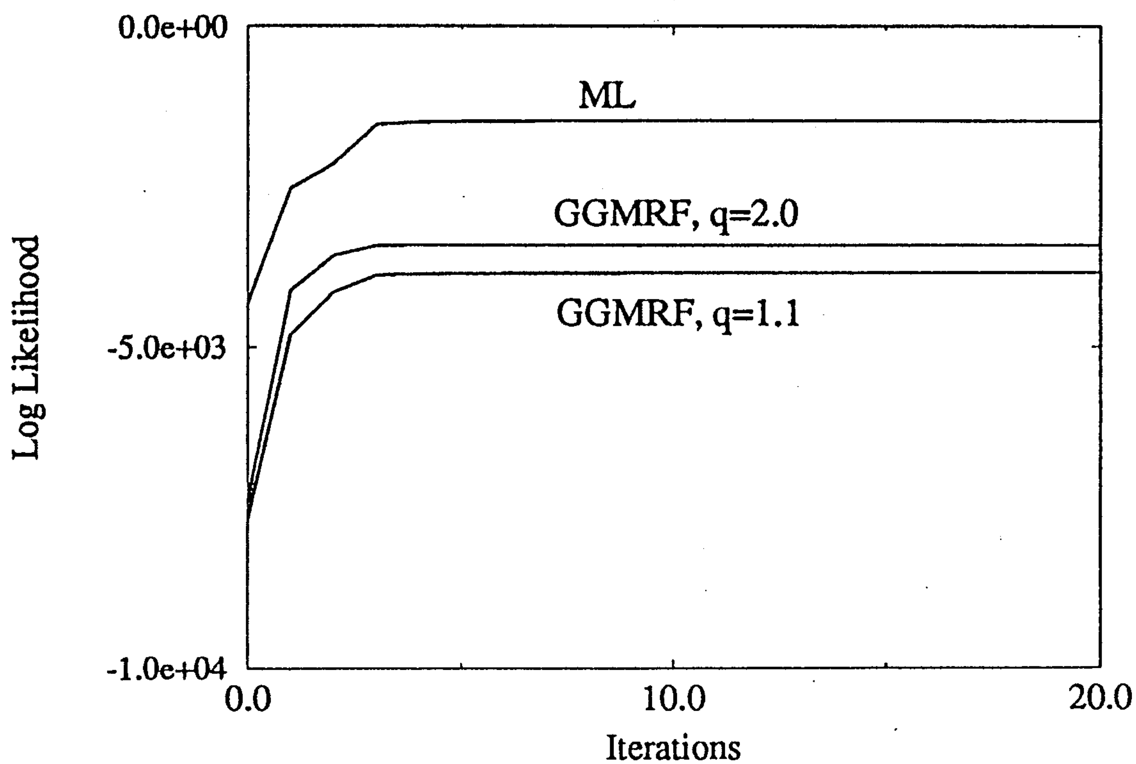


Figure 6: Convergence of ML and MAP estimates using ICD/Newton-Raphson updates. The *a posteriori* likelihood function values are plotted as a function of full iterations. Top plot is ML, center is Gaussian prior, bottom is GGMRF with $q = 1.1$.