Mad: Linear Estimation.

$$y = Ax + \nu \qquad \text{Data: } (y, A, S)$$

covar matr of ν

$\operatorname{Var} \nu$

Assumed: x – not known.

$$\hat{x} = Ry = (A^* S^{-1} A)^{-1} A^* S^{-1} y$$

$$\begin{array}{ccccc}
(y_1, A_1, S_1) \rightarrow (T_1, z_1) & \searrow & & & \\
\vdots \qquad\qquad \vdots & & \oplus \rightarrow (T, z) & \rightarrow & \hat{x} = T^{-1} z \\
(y_n, A_n, S_n) \rightarrow (T_n, z_n) & \nearrow & & & Q = T^{-1}
\end{array}$$

$$T = \sum_{i=1}^{n} T_i \qquad z = \sum_{i=1}^{n} z_i$$

$$T_i = A_i^T S_i^{-1} A_i, \quad z_i = A_i^T S_i^{-1} y_i$$

$Q = T^{-1}$: Precision

$$\underline{E} \|\hat{x} - x\|^2 = \operatorname{tr} Q \qquad \operatorname{tr} Q = \sum_{i=1}^{n} Q_{ii}$$

$$E(\hat{x}_i - x_i)^2 = Q_{ii}$$

Estimation with prior information.

$$y = Ax + \nu$$

ν – rand. vec $\quad E\nu = 0 \quad \operatorname{Var} \nu = S$

A – Linear transf: $A: \mathcal{D} \to \mathcal{R}$ (matrix)

x – unknown

Prior info about x. in probabilistic

form.

Treat x - as random vector

with $Ex = x_0$

$\mathrm{Var}\, x = F$ - covar matrix.

$F > 0$ - posotive definite

$\Rightarrow$ invertible.

Prior info (x_0, F)

Raw Data (y, A, S)

$$\hat{x} = Ry + r \qquad R: \mathcal{R} \to \mathcal{D}, \quad r \in \mathcal{D}$$

$$R, r = ?$$

$$E_{\nu, x} \|\hat{x} - x\|^2 = \underbrace{H(R, r)}\ \sim \min_{R, r}$$

$H_{min} = \mathrm{tr}\, Q \qquad Q$ - precision matrix

$$Q = \left(A^T S^{-1} A + F^{-1}\right)^{-1}$$

$$\hat{x} = Q\left(A^T S^{-1} y + F^{-1} x_0\right)$$

$$E(\hat{x}_i - x_i)^2 = Q_{ii}$$

Vanishing Prior info.

$F \to \infty$ when $F^{-1} \to 0$

$$Q = \left(A^T S^{-1} A + \underbrace{F^{-1}}_{\to 0}\right)^{-1} \to \left(A^T S^{-1} A\right)^{-1}$$

same as for unknown x

$$\hat{x} = Q\left(A^T S^{-1} y + F^{-1} x_0\right) \to \left(A^T S^{-1} A\right)^{-1} A^T S^{-1} y$$

same unknown x

$$\hat{x} = Q(A^T S^{-1} y + \underbrace{F^{-1}}_{\to 0} x_0) \to \underbrace{(A^T S^{-1} A)^{-1} A^T S^{-1} y}$$

$\underbrace{\quad\quad}_{\to 0}$

Prior info as an additional measurement.

$$\begin{cases} y = Ax + \nu & \text{original observation} \\ x_0 = Ix + \mu & \text{hypothetical} \ -\text{"}- \end{cases}$$

$E\mu = 0$, $Var\,\mu = F$

instead of prior info (x_0, F)

$$\begin{bmatrix} y \\ x_0 \end{bmatrix} = \begin{bmatrix} A \\ I \end{bmatrix} x + \begin{bmatrix} \nu \\ \mu \end{bmatrix} \longrightarrow \text{Raw info}$$

$$\left(\begin{bmatrix} y \\ x_0 \end{bmatrix}, \begin{bmatrix} A \\ I \end{bmatrix}, \begin{bmatrix} S & 0 \\ 0 & F \end{bmatrix} \right)$$

$$\hat{x} = \underbrace{\left(\begin{bmatrix} A^T & I \end{bmatrix} \underbrace{\begin{bmatrix} S & 0 \\ 0 & F \end{bmatrix}^{-1}}_{\begin{bmatrix} S^{-1} & 0 \\ 0 & F^{-1} \end{bmatrix}} \begin{bmatrix} A \\ I \end{bmatrix} \right)^{-1}}_{Q =} \begin{bmatrix} A^T & I \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & F \end{bmatrix}^{-1} \begin{bmatrix} y \\ x_0 \end{bmatrix} =$$

$$= (A^T S^{-1} A + I F^{-1} I)^{-1} (A^T S^{-1} y + I F^{-1} x_0) =$$

$$= \underbrace{\underbrace{(A^T S^{-1} A + F^{-1})^{-1}} (A^T S^{-1} y + F^{-1} x_0)}$$

same as for est w. prior info

$$Q = (A^T S^{-1} A + F^{-1})^{-1}$$

Prior info can be represented with can info

can. form of prior info.

Prior info
with can info

can. form of prior info.

$$(x_0, F) \longrightarrow (T_0, b_0) = (F^{-1}, F^{-1}x_0)$$

can be treated as observation

$$(x_0, F) \longrightarrow (x_0, I, F) \longrightarrow (T_0, b_0)$$

$$T_0 = I^T F^{-1} I = F^{-1}$$
$$b_0 = I^T F^{-1} x_0 = F^{-1} x_0$$

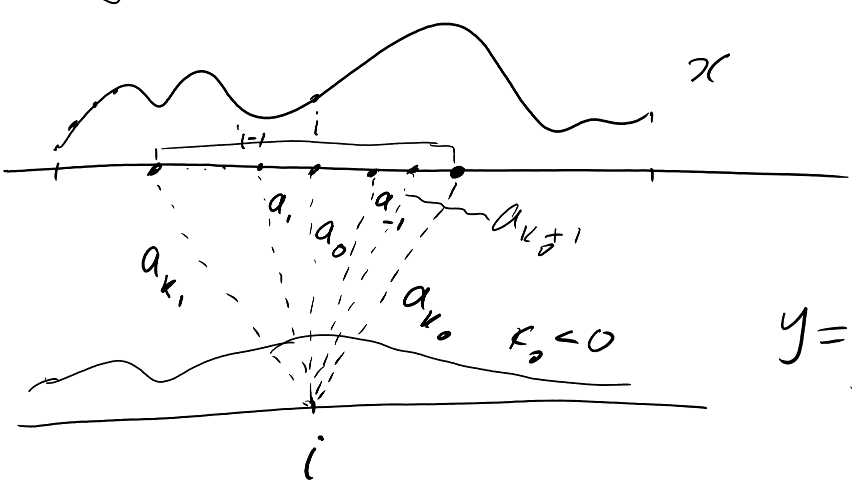
$$\begin{array}{lll}
\text{prior info } (x_0, F_0) \longrightarrow (T_0, b_0) & & \\
\text{Raw Data} \left\{ \begin{array}{l} (y_1, A_1, S_1) \longrightarrow (T_1, b_1) \\ \quad \vdots \qquad\qquad\quad \vdots \\ (y_n, A_n, S_n) \longrightarrow (T_n, b_n) \end{array} \right. & \longrightarrow \oplus \longrightarrow (T, b) \longrightarrow \hat{x}, Q &
\end{array}$$

$$T = \sum_{i=0}^{n} T_i$$
$$b = \sum_{i=0}^{n} b_i$$

$$Q = T^{-1}$$
$$\hat{x} = T^{-1} b$$

---

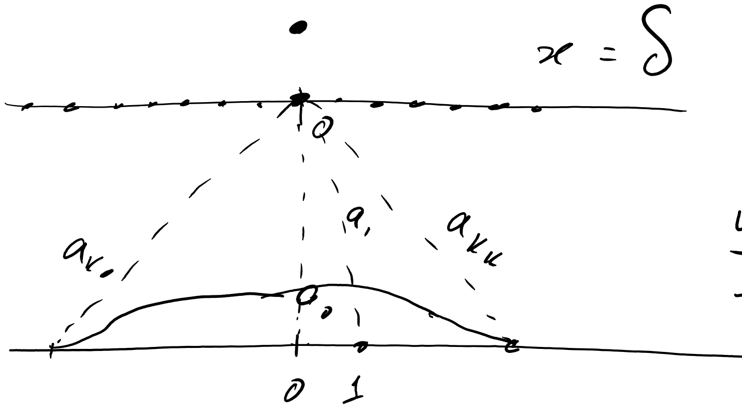
for HW #6



$$y = \underline{\underline{A}} x$$

$$y_i = \sum_{k=k_0}^{k_1} a_k x_{i-k} \quad \text{– convolution.}$$

$$y_i = \sum_{k=k_0} a_k x_{i-k}$$ – convolution

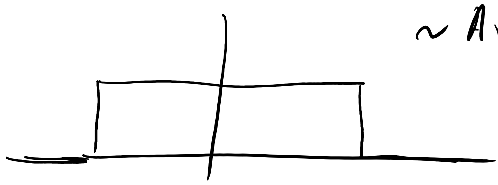


$x = \delta$: $x_i = \begin{cases} 1 & i = 0 \\ 0 & i \neq 0 \end{cases}$

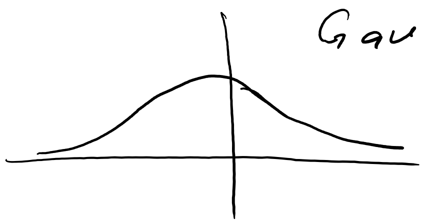
$$\underline{y_i} = \sum_{k=k_0}^{k_1} a_k \delta_{i-k} =$$

$$= \underline{a_i}$$ – Point Spread function (PSF)

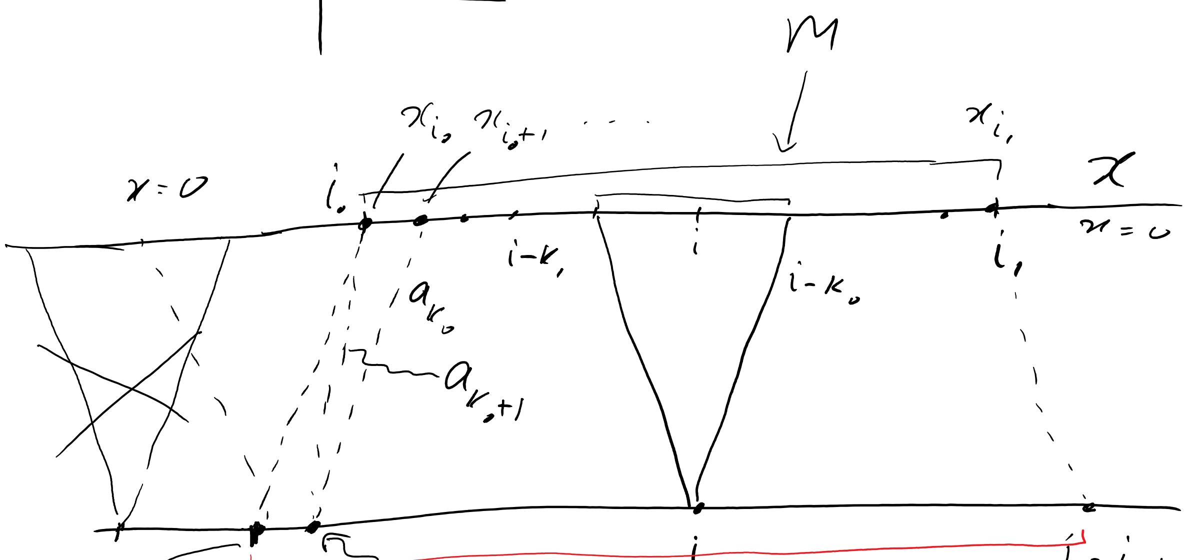
PSFs:



~ Averaging



Gauss



$j_1 = i_1 + k_1$

$j_0 = i_0 + k_0$

$$j_1 - j_0 = i_1 + k_1 - (i_0 + k_0) = \underbrace{i_1 - i_0}_{= m} + \underbrace{k_1 - k_0}_{\text{width of the sliding window}}$$

$$y_{j_0} = x_{i_0} a_{k_0} \qquad y_{j_0+1} = a_{k_0+1} x_{i_0} + a_{k_0} x_{i_{0}+1}$$

PSF

$[y]$ $[a\ 0\ - - - -\ 0]$

Sliding window

$$
\begin{array}{c}
n = \\ m + (-k_0) + k_1 = \\ = m + k_1 - k_0
\end{array}
\begin{bmatrix} y_{j_0} \\ y_{j_0+1} \\ \vdots \\ \vdots \\ y_{j_1} \end{bmatrix}
=
\begin{bmatrix}
a_{k_0} & 0 & \cdots & & 0 \\
a_{k_0+1} & a_{k_0} & 0 & \cdots & 0 \\
a_{k_0+2} & a_{k_0+1} & a_{k_0} & & \\
\vdots & & & \ddots & \\
a_0 & & & & \\
a_1 & & & & a_{k_0} \\
\vdots & & & & \\
a_{k_1} & & & & a_0 \\
0 & & & & \\
\vdots & & & & a_{k_1} \\
0 & & & &
\end{bmatrix}
\begin{bmatrix} x_{i_0} \\ x_{i_0+1} \\ \vdots \\ x_{i_1} \end{bmatrix}
\; m
$$

$$y = Ax + \nu$$

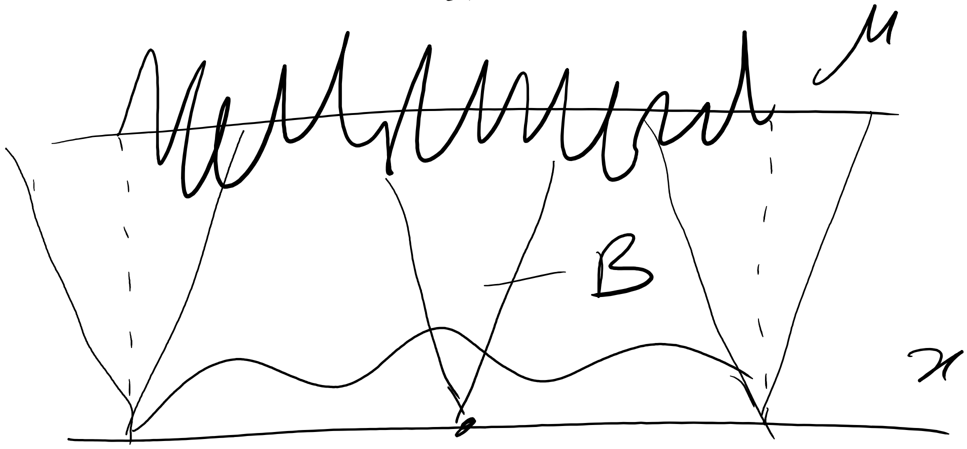
Generating random x with $Ex = 0$

$\quad Var\, x = F$,

Take μ : $E\mu = 0$ $\quad Var\, \mu = I$

e.g. $\mu_i \sim N(0, 1)$

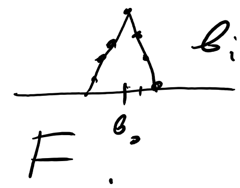
normal mean var.



white noize

$\underline{\mu_i \text{ - i.i.d.}}$

Take $$\mathcal{B} = \begin{bmatrix} b_0 & b_1 & b_2 & \cdots & \\ b_1 & \ddots & & \ddots & b_2 \\ b_2 & \ddots & \ddots & & b_1 \\ & \ddots & b_2 & b_1 & b_0 \end{bmatrix}$$ Toeplitz matrix.



$x = \mathcal{B}\mu$

$$\operatorname{Var} x = \mathcal{B} \cdot \underbrace{\operatorname{var}\mu}_{I} \cdot \mathcal{B}^T = \mathcal{B}\mathcal{B}^T = F.$$

---

main steps of HW # 6.

* generate random x
  - generate μ (white noise
  - create Toeplitz matrix $\mathcal{B}$
  - $x = \mathcal{B}\mu$
  - $F = \mathcal{B}\mathcal{B}^T$ – know.
* Simulate measurement.
  - generate A
  - $y = Ax + \nu$,
  - get raw info
    (y, A, S)

    $S = \operatorname{Var}\nu$ $\quad \nu_i$ – i.i.d

    $\nu_i \sim (0, \sigma^2)$

    $S = \sigma^2 I \Rightarrow \underline{S^{-1} = \frac{1}{\sigma^2} I}$
  - transform to canonical
    (T, v)
* Construct estimate
  - $\hat{x}$ without prior info
  - $\hat{x}$ with prior info

– $\hat{x}$ with prior info

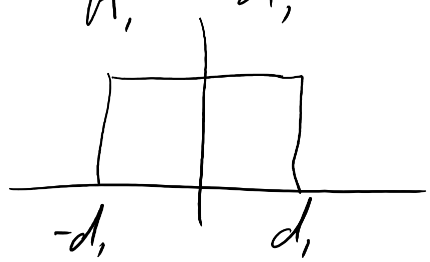
$$(0, F) \to (T_0, v_0) = (T_0, 0)$$

$$T_0 = F^{-1} \quad v_0 = F_0^{-1} x_0 = 0$$

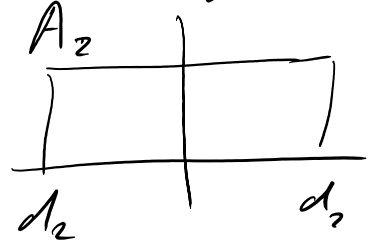
for 1 or <u>many measurements</u>.

A – takes different width.

$A_1 \leftrightarrow d_1$



A_2



$10 \le d_i \le 20$

Transition from prior to posterior information

prior: $x \sim (x_0, F_0)$

mean, cov. matr.

$y_1 = A_1 x + v_1$ raw inf (y_1, A_1, S_1)

prior (x_0, F_0)

Raw $(y_1, A_1, S_1) \to \circledast \to (x_1, F_1)$ – posterior information.

$\hat{x}$, Q – new (updated) info about x.

$(y_2, A_2, S_2) \to \circledast \to (x_2, F_2)$

⋮

(x_{k-1}, F_{k-1})

$(y_k\, A_k\, S_k) \to \circledast \to (x_k, F_k)$

$$(y_k\, A_k\, S_k) \longrightarrow \circledast \longrightarrow (x_k, F_k)$$

$$F_k = \left(A_k^T S_k^{-1} A_k + F_{k-1}^{-1}\right)^{-1}$$

$$x_k = F_k \left(A_k^T S_k^{-1} y_k + F_{k-1}^{-1} x_{k-1}\right)$$

Convenient for Big Data Streams.