Review: Weighted LS

Model: $y = Hx$

Meas: $\tilde{y} = Hx + v$

$e = \tilde{y} - y$ W: diag. weights

$\min_x \frac{1}{2} e^T e = (\tilde{y} - Hx)^T (\tilde{y} - Hx) \Rightarrow$ Soln: $\hat{x} = (H^T W H)^{-1} H^T W \tilde{y}$

BATCH LS: $H_1 \in \mathbb{R}^{m_1 \times n}$ $H_2 \in \mathbb{R}^{m_2 \times n}$

2 sets of meas:

$\tilde{y}_1 = H_1 \boxed{x} + v_1$ } init. set of meas

$\tilde{y}_2 = H_2 \boxed{x} + v_2$ } add. set of meas

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} x + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

W_1: diag $W_1 \succ 0$

W_2: diag $W_2 \succeq 0$

if we've already solved for $\hat{x}_1 = (H_1^T W_1 H_1)^{-1} H_1^T W_1 \tilde{y}_1$

can we leverage $\hat{x}_1$ to compute new est. w/ all data.

in particular:

H_1 : tall $\rightarrow$ lot of data pts.

H_2 : fat $\rightarrow$ just adding a few rows / data pts.

want to compute new estimate: $\hat{x}_2$

$$\hat{x}_2 = \left(\begin{bmatrix} H_1^T & H_2 \end{bmatrix} \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \right)^{-1} \begin{bmatrix} H_1^T & H_2^T \end{bmatrix} \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix}$$

$$= \overset{\star}{\Big[} \underbrace{H_1^T W_1 H_1 + H_2^T W_2 H_2}_{\text{need to invert.}} \Big]^{-1} \left[H_1^T W_1 \tilde{y}_1 + H_2^T W_2 \tilde{y}_2 \right]$$

already computed $(H_1^T W_1 H_1)^{-1}$

can we leverage to make inverting $\star$ easier?

if H_2 fat: (just a few new data pts)

$$\Big[\underbrace{H_1^T W_1 H_1}_{n \times n} + H_2^T \underbrace{W_2}_{m_2 \times m_2} H_2 \Big]^{-1}$$

$$\Big[H_1^T W_1 H_1 \Big] + \underbrace{\Big[H_2^T \Big] \Big[\underset{m_2 \times m_2}{} \Big] \Big[\quad H_2 \quad \Big]}_{\substack{\text{adding a low} \\ \text{rank piece}}}$$

Wood Bury Matrix Identity.

$$(A + UCV)^{-1} = A^{-1} - A^{-1} U (C^{-1} + V A^{-1} U)^{-1} V A^{-1}$$

$$\left[\quad\right] + [u][\;][\;]^{-1} = [A^{-1}] - [A^{-1}u]\left[C^{-1} + VA^{-1}u\right]^{-1}[VA^{-1}] \leftarrow$$

one $n \times n$ inverse — for 2 $m_2 \times m_2$ and some matrix multiplications

order n^3 ($n^{2.37}$)

if $n \gg m_2$

$\Rightarrow$ on order n^2 (—)

$$\hat{x}_2 = (H_1^T W_1 H_1 + H_2^T W_2 H_2)^{-1}\left[H_1^T W_1 \tilde{y}_1 + H_2^T W_2 \tilde{y}_2\right]$$

Let $P_1 = (H_1^T W_1 H_1)^{-1}$, $P_2 = (H_1^T W_1 H_1 + H_2^T W_2 H_2)^{-1}$

Woodbury:

$$P_2 = P_1 - \underbrace{P_1 H_2^T (W_2^{-1} + H_2 P_1 H_2^T)^{-1}}_{K_2} H_2 P_1 \leftarrow$$

$$\boxed{\text{Let } K_2 := P_1 H_2^T (W_2^{-1} + H_2 P_1 H_2^T)^{-1} \overset{*}{=} P_2 H_2^T W_2}$$

not obvs.

Proof of $*$. $P_2 H_2^T W_2 = P_1 H_2^T W_2 - \underbrace{P_1 H_2^T (W_2^{-1} + H_2 P_1 H_2^T)^{-1}} H_2 P_1 H_2^T W_2$

$$= P_1 H_2^T \left(W_2 - (W_2^{-1} + H_2 P_1 H_2^T)^{-1} H_2 P_1 H_2^T W_2\right)$$

$$= \underbrace{P_1 H_2^T (W_2^{-1} + H_2 P_1 H_2^T)^{-1}}_{K_2} \left[\underbrace{(\cancel{W_2^{-1}} + \cancel{H_2 P_1 H_2^T}) - \cancel{H_2 P_1 H_2^T}}_{I}\right] \cancel{W_2}$$

$$\boxed{K_2 = P_1 H_2^T (W_2^{-1} + H_2 P_1 H_2^T)^{-1}}$$

like KF gain

$$\boxed{P_2 = (I - K_2 H_2) P_1}$$

like covariance update

$$\hat{x}_2 = \underset{\sim}{P_2} H_1^T W_1 \tilde{y}_1 + \underline{P_2 H_2^T W_2} \tilde{y}_2$$

$$= (I - K_2 H_2) \underbrace{P_1 H_1^T W_1 \tilde{y}_1}_{\hat{x}_1} + K_2 \tilde{y}_2$$

$$\boxed{\hat{x}_2 = \underline{\hat{x}_1} + \underbrace{K_2}_{\text{update gain}} (\underline{\tilde{y}_2} - \underline{H_2 \hat{x}_1})} \Leftarrow$$

like state update

if $\hat{x}_1$ is correct what will $\tilde{y}_2$ be? ie. prediction of $\underline{\tilde{y}_2}$

modification to the estimate

SEQUENTIAL LS:

have $\hat{x}_k$, K_k, P_k

$$K_{k+1} = P_k H_{k+1}^T \left(W_{k+1}^{-1} + H_{k+1} P_k H_{k+1}^T \right)^{-1}$$

$$P_{k+1} = (I - K_{k+1} H_{k+1}) P_k$$

$$\hat{x}_{k+1} = \hat{x}_k + K_{k+1} (\underline{y_{k+1}} - \underline{H_{k+1}} \underline{\underline{\hat{x}_k}})$$

x KF $x^+ = Ax + Bu.$

Nonlinear LS: (Gauss-Newton ?)

Levenberg-Marquadt (another version)

Model: $y = f(x)$

Meas $\tilde{y} = f(x) + v$

Iterative linearization process

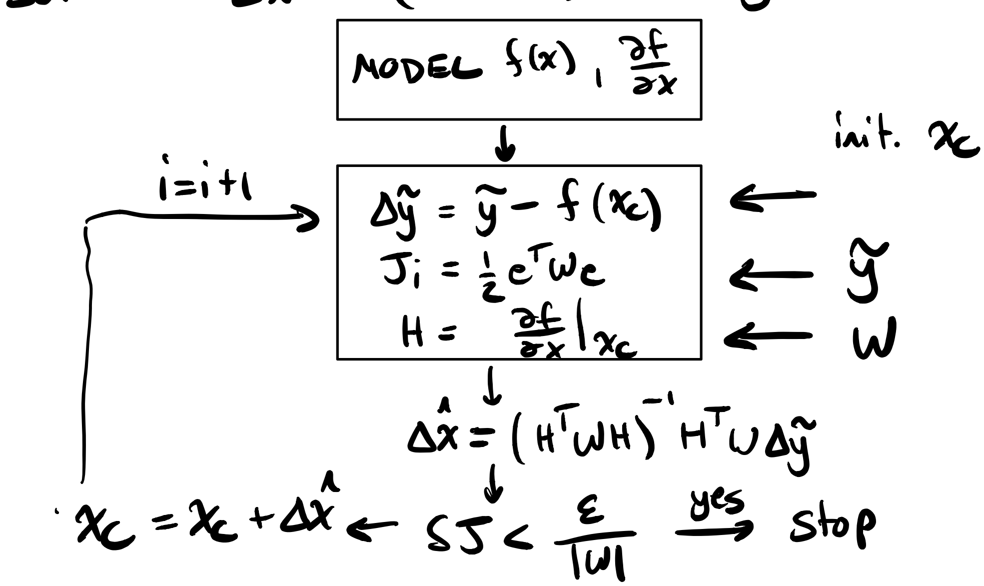
$$x = \underline{x_c} + \underline{\Delta x}$$

$$f(x) = f(x_c + \Delta x) \approx f(x_c) + \underbrace{\left.\frac{\partial f}{\partial x}\right|_{x_c}}_{H} \Delta x$$

$$e = \tilde{y} - f(x) = \underbrace{\tilde{y} - f(x_c)}_{\Delta\tilde{y}} - H\Delta x$$

$$\min_{\Delta x} \ \frac{1}{2} e^T W e = \frac{1}{2}(\Delta\tilde{y} - H\Delta x)^T(\Delta\tilde{y} - H\Delta x)$$

solution: $\Delta\hat{x} = (H^T W H)^{-1} H^T W \Delta\tilde{y}$



MINIMUM VARIANCE ESTIMATION

$\tilde{y} = \underline{H}x + \underline{v}$ give v more structure $v \sim N(0, R)$

$R = R^T > 0$

what is an estimator...

$\hat{x}(\tilde{y}) \Leftarrow$ function of meas

Assume some class of functions for $\hat{x}(\cdot)$

find best $\hat{x}(\tilde{y})$ w/in that class of functions

Estimator Bias "is the expected value of the estimated parameters consistent w/ true parameters"

BIAS : $E[\hat{x}(\tilde{y}) - x]$

UNBIASED $\Rightarrow E[\hat{x}(\tilde{y})] = x$

Linear Estimator (assume)

$\hat{x} = M\tilde{y} + n$ picking $\hat{x}(\cdot)$ is now picking M, n

What conds do we need on M, n for an unbiased estimator?

$$E[\hat{x}] = E[M\tilde{y} + n] = E[MHx + Mv + n]$$

$$= \underline{MH}x + n + M\cancel{E[v]}^{0} \underset{\text{want.}}{=} x$$

$\Rightarrow$ imposes constraints on M & n...

$$\left.\begin{array}{l} \bullet\ n = 0 \\ \bullet\ MH = I \end{array}\right\} \rightarrow \text{guarantee that our unbiased}$$

NOTE: $M = H^{-1}$?

H is tall. can't invert H...

if $MH = I$ and $ZH = 0 \Rightarrow (M+Z)H = I$

since H is tall multple M's can work.

- $M = (H^T H)^{-1} H^T \qquad (H^T H)^{-1} H^T \times H = I$
- $M = (H^T W H)^{-1} H^T W \qquad (H^T W H)^{-1} H^T W \times H = I$
- take any C s.t. CH is invertible...

$$M = (CH)^{-1} C \Rightarrow (CH)^{-1} C \times H = I$$

$$[-C-][H]$$

unbiased $\Rightarrow$ gave us constraints..

now choose $M, (n)$ to minimize variance

$$J = \frac{1}{2} E\left[(\hat{x} - x)^T (\hat{x} - x)\right]$$

$$= \frac{1}{2} E \sum_i (\hat{x}_i - x_i)^2$$

$$= \mathrm{Tr}\,\frac{1}{2}\,E\left[(\hat{x}-x)^T(\hat{x}-x)\right] \qquad \mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

$$= \frac{1}{2}\,\mathrm{Tr}\,E\left[\underbrace{(\hat{x}-x)(\hat{x}-x)^T}_{\text{min trace of covariance of } \hat{x}}\right]$$

$$\hat{x} = M\tilde{y} + \cancel{n}^{\,0} \qquad \tilde{y} = Hx + v$$

$$= \underbrace{MH}_{\text{needs to be } I}x + M\underline{v} \qquad v \sim N(0, R)$$

$$= x + Mv$$

$$J = \frac{1}{2}\,\mathrm{Tr}\,E\left[Mvv^TM^T\right] = \frac{1}{2}\,\mathrm{Tr}\,(MRM^T)$$

optimization:

$$\left.\begin{aligned} &\min_M \quad J = \frac{1}{2}\mathrm{Tr}(MRM^T) \\ &\text{s.t.} \quad MH = I \end{aligned}\right| \rightarrow \quad \underline{\text{HOMEWORK}}$$

$$M = (H^TR^{-1}H)^{-1}H^TR^{-1}$$

$$\Rightarrow \quad \hat{x} = M\tilde{y} = (H^TR^{-1}H)^{-1}H^TR^{-1}\tilde{y}$$

this is weighted LS with $W = R^{-1}$

R^{-1} not diag , $R^{-1} > 0$

How do I weight my measurements properly to account for noise v?

deps on structure of v $\quad v \sim N(0, R)$

$\Rightarrow$ use $W = R^{-1}$ $\leftarrow$ meas. w/ more noise get weighted less.

<u>Solving</u> HW1

$$L(M, V) = \frac{1}{2} \mathrm{Tr}(\underline{M} R \underline{M}^T) + \mathrm{Tr}(\underline{V}^T (MH - I))$$

for matrices

$$\langle A, B \rangle = \sum_{ij} A_{ij} B_{ij} = \mathrm{Tr}(A^T B)$$

$$\frac{\partial L}{\partial M} = 0 \qquad \frac{\partial L}{\partial V} = 0$$

$$\frac{\partial}{\partial M} \frac{1}{2} \mathrm{Tr}(M R M^T)$$

$$f(M) = \mathrm{Tr}(M R M^T)$$

$$\Delta f = \mathrm{Tr}(\Delta M R M^T) + \mathrm{Tr}(\underbrace{M R \Delta M^T}_{\text{transpose}})$$

$$= \mathrm{Tr}(\Delta M R M^T) + \mathrm{Tr}(\Delta M R^T M)$$

$$= \mathrm{Tr}(\Delta M (R + R^T) M)$$

$$= \langle \Delta M^T, \underline{(R+R^T) M} \rangle = \langle (R+R^T) M, \Delta M^T \rangle$$

$$f(x) = \langle C, x \rangle \Rightarrow \frac{\partial f}{\partial x} = C$$

$$\frac{\partial f}{\partial M} = M^T (R + R^T)$$