Minimum Variance Estimation

Goal: ① Connect previous results to probability theory

② Understand the weight matrix W in weighted least squares

From before:

$$\underset{(m\times 1)}{\tilde{\underline{y}}} = \underset{(m\times n)}{H}\,\underset{(n\times 1)}{\underline{x}} + \underset{(m\times 1)}{\underline{v}} \qquad (2.1)$$

Estimate $\underline{x}$:

$$\underset{(n\times 1)}{\hat{\underline{x}}} = \underset{(n\times m)}{M}\,\underset{(m\times 1)}{\tilde{\underline{y}}} + \underset{(n\times 1)}{\underline{n}}$$

What are M and $\underline{n}$?

The optimal choice of M and $\underline{n}$ would minimize each individual estimate:

$$J_i = \frac{1}{2}\underbrace{E\left\{(\hat{x}_i - x_i)^2\right\}}_{\text{"expected value"}}, \quad i = 1, 2 \ldots n \qquad (2.3)$$

"expected value"

= average or mean value

If there is no noise $(\underline{v} = 0)$, then:

$$\tilde{\underline{y}} \equiv \underline{y} = H\underline{x}$$

Thus, we know the following must be true:

$$\underline{x} = MH\underline{x} + \underline{n}$$

So M & $\underline{n}$ must satisfy:

$$\underline{n} = 0$$

$$MH = I, \quad H^T M^T = I \quad \leftarrow \text{"constraint"}$$

We know our optimal estimator must look like:

$$\hat{\underline{x}} = M\tilde{\underline{y}} \tag{2.8}$$

To facilitate subsequent manipulations:

$$M = \begin{bmatrix} - m_1 - \\ - m_2 - \\ \vdots \\ - m_n - \end{bmatrix} \quad I = \begin{bmatrix} - I_1^r - \\ - I_2^r - \\ \vdots \\ - I_n^r - \end{bmatrix} = \begin{bmatrix} | & | & & | \\ I_1^c & I_2^c & \cdots & I_n^c \\ | & | & & | \end{bmatrix}$$

note: $I_i^r = (I_i^c)^T$

Have students do this:

Now rewrite constraint as:

rows of M

$$M_i H = I_i^r \qquad (2.12)$$

Now i^{th} row of $\underline{\hat{x}} = M\tilde{y}$:

$$\hat{x}_i = M_i \tilde{y}$$

$\hat{x}_i$ depends only on i^{th} row of M.
This is helpful: we can optimize each $\hat{x}_i$ independently!

We can now rewrite 2.3:

$$J_i = \frac{1}{2} E \left\{ \left(M_i \tilde{y} - x_i \right)^2 \right\}$$

Plug in 2.1: (the eqn for $\tilde{y}$)

$$J_i = \frac{1}{2} E \left\{ \left(M_i H \underline{x} + M_i v - x_i \right)^2 \right\}$$

Bring in 2.12:

$$J_i = \frac{1}{2} E \left\{ \left(\underbrace{I_i^r \underline{x}}_{= x_i} + M_i \underline{v} - x_i \right)^2 \right\}$$

Result: the x_i's cancel, leaving:

$$J_i = \frac{1}{2} E \left\{ \underbrace{\underset{[1 \times m]}{M_i} \underbrace{(\underset{[m\times1]}{\underline{v}}\, \underset{[1\times m]}{\underline{v}^T})}_{m \times m} \underset{m \times 1}{M_i^T}}_{1 \times 1} \right\}$$

M is not a random variable, only $\underline{v}\,\underline{v}^T$ is, so rewrite:

$$J_i = \frac{1}{2} M_i \underbrace{E\{\underline{v}\,\underline{v}^T\}} M_i^T$$

"covariance matrix"
(of measurements)
= "R"

if all measurements are independent, then:

$$R = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_m^2 \end{bmatrix}$$

variance of measurement m

So we have:

$$J_i = \frac{1}{2} m_i R m_i^T$$

Minimize J_i subject to constraint:

$$m_i H = I_i^r \quad \text{or} \quad H^T m_i^T = I_i^c$$

Solve using method of Lagrange Multipliers:

$$J_i = \frac{1}{2} m_i R m_i^T + \underline{\lambda}_i^T \underbrace{\left(I_i^c - H^T m_i^T \right)}_{\text{constraint}}$$

① $$\nabla_{m_i^T} J_i = R m_i^T - H \underline{\lambda}_i = 0 \qquad (2.23)$$

② $$\nabla_{\lambda_i} J_i = I_i^c - H^T m_i^T = 0 \qquad (2.24)$$

$$I_i^r - m_i H = 0$$

same

2 eqns, 2 unknowns $(\underline{\lambda}_i \;\&\; m_i)$

Solve for m_i:

$$m_i = I_i^r \left(H^T R^{-1} H \right)^{-1} H^T R^{-1}$$

$$m = \left(H^T R^{-1} H \right)^{-1} H^T R^{-1}$$

So we have:

$$\boxed{\hat{\underline{x}} = (H^T R^{-1} H)^{-1} H^T R^{-1} \tilde{\underline{y}}}$$

Gauss Markov Theorem

Same as weighted LLS result, but now we know that $W = R$ is optimal.