LEAST SQUARES (CONT):

Model: $y = Hx$

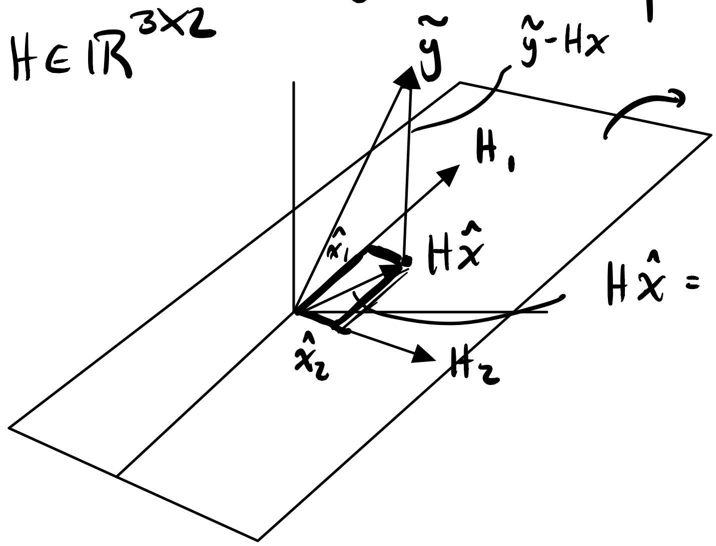
Meas: $\tilde{y} = Hx + v$

$$\min_x |\tilde{y} - Hx|^2 = (\tilde{y} - Hx)^T(\tilde{y} - Hx) \Leftarrow$$

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

PICTURE $\tilde{y} = Hx = \begin{bmatrix} H_1 & H_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Leftarrow$

$H \in \mathbb{R}^{3 \times 2}$



$$H\hat{x} = \underbrace{H(H^T H)^{-1} H^T}_{\text{proj}_{\mathcal{R}(H)}} \tilde{y}$$

$\text{proj}_{\mathcal{R}(H)} \tilde{y}$

Example: System Identification

Discrete LTI SYS

$$z^+ = \Phi z + \Gamma u$$

want to find Φ, Γ

$z \in \mathbb{R}^k$

$\Phi \in \mathbb{R}^{k \times k}$

$\Gamma \in \mathbb{R}^{k \times l}$

$u \in \mathbb{R}^l$

FROM DATA: $\underbrace{[z_0\ z_1 \cdots z_T]}_{\text{known}}$, $\underbrace{[u_0 \cdots u_{T-1}]}_{\text{known}}$

$$\underbrace{\tilde{y}}_{\text{known}} = \underbrace{H}_{\text{known}}\ \underbrace{x}_{\text{unknown}}$$

LS: what is x? x is Φ, Γ

what is H? H from z_t, u_t

what is $\tilde{y}$? z_t (at next time step)

Model

$$z_{t+1} = \Phi z_t + \Gamma u_t$$

$$[z_1 \cdots z_T] = \Phi [z_0 \cdots z_{T-1}] + \Gamma [u_0 \cdots u_{T-1}]$$

$$= [\Phi\ \ \Gamma] \begin{bmatrix} z_0 \cdots z_{T-1} \\ u_0 \cdots u_{T-1} \end{bmatrix}$$

transpose...

$$T\ \underbrace{\overset{k}{\begin{bmatrix} z_1^T \\ \vdots \\ z_T^T \end{bmatrix}}}_{\tilde{y}} = T\ \underbrace{\overset{k+l}{\begin{bmatrix} z_0^T & u_0^T \\ \vdots & \vdots \\ z_{T-1}^T & u_{T-1}^T \end{bmatrix}}}_{H}\ {\scriptstyle k+l}\ \underbrace{\overset{k}{\begin{bmatrix} \Phi \\ \Gamma \end{bmatrix}}}_{x}$$

k individual LS problems

$$\underline{\hat{x}} = \begin{bmatrix} \hat{\Phi} \\ \hat{\Gamma} \end{bmatrix} = (H^T H)^{-1} H^T \underline{\tilde{y}}$$

Variations on Least Squares:

Weighted LS:

$e = \tilde{y} - Hx$

$$\min_x J = e^T e = (\tilde{y} - Hx)^T(\tilde{y} - Hx) = \sum_i e_i^2$$

if we trust some measurements more than others... add a weighting matrix...

W : diagonal $\quad W \succ 0 \quad W \in \mathbb{R}^{m \times m}$

trust meas i : w_{ii} larger $\quad w_{ii} > 0$

don't trust meas i : w_{ii} smaller

Modified opt problem:

$$\min_x J = \frac{1}{2} e^T W e = \frac{1}{2}(\tilde{y} - Hx)^T W (\tilde{y} - Hx) = \frac{1}{2}\sum_i e_i^2 w_i$$

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

Constrained Least Squares

$$\tilde{y}_1 = H_1 x + e_1 \Leftarrow \text{uncertain meas.}$$

$$\tilde{y}_2 = H_2 x \Leftarrow \text{certain meas}$$

- treating as a set of constraints on x.

$$\tilde{y} = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} x + \begin{bmatrix} e_1 \\ 0 \end{bmatrix}$$

$H_1 \in \mathbb{R}^{m_1 \times n}$, $H_2 \in \mathbb{R}^{m_2 \times n}$ need $m_2 < n$, $m_1 >> n$

(H_2 needs to have a non trivial nullspace for this problem to be well-posed)

Ex. meas position of robot

$\tilde{y}_1 = H_1 x + e_1 \rightarrow$ position of arm in degrees of freedom

$\tilde{y}_2 = H_2 x \rightarrow$ constraints on position enforced by the joints

$$\min_x J = \frac{1}{2} e_1^T W_1 e_1 = \frac{1}{2}(\tilde{y}_1 - H_1 x)^T W_1 (\tilde{y}_1 - H_1 x)$$

objective

$$\text{s.t. } \tilde{y}_2 = H_2 x \rightarrow \text{constraints.}$$

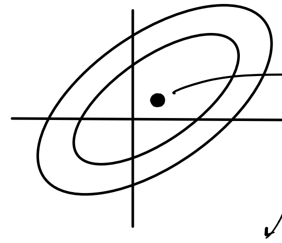
Before: optimality conds $\frac{\partial J}{\partial x} = 0$

Now: opt conds $\frac{\partial L}{\partial x} = 0 \quad \frac{\partial L}{\partial v} = 0$

$$L(x,v) = \frac{1}{2}(\tilde{y}_1 - H_1 x)^T W_1 (\tilde{y}_1 - H_1 x) + v^T(\tilde{y}_2 - H_2 x)$$

Method of Lagrange Multipliers:

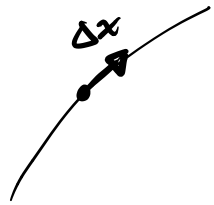
Unconstrained: $\min_x J(x)$



Constrained: $\min_x J(x)$
s.t. $g(x) = 0$

$\frac{\partial f}{\partial x} \perp$ level sets of f

$$\Delta f = \frac{\partial f}{\partial x} \Delta x$$



for level set

$$0 = \frac{\partial f}{\partial x} \underbrace{\Delta x}_{\text{along level set}}$$

Optimality cond:

$$\frac{\partial J}{\partial x} + v \frac{\partial g}{\partial x} = 0$$

$$\left[\frac{\partial J}{\partial x} + v^T \frac{\partial g}{\partial x} = 0 \right]$$

What function L should we minimize so that $\frac{\partial L}{\partial x} = \frac{\partial J}{\partial x} + v^T \frac{\partial g}{\partial x}$?

gradient of J is a linear comb of rows of $\frac{\partial g}{\partial x}$

$$L = J(x) + v^T g(x) \Leftarrow \text{Lagrangian}$$

$$\frac{\partial L}{\partial x} = \frac{\partial J}{\partial x} + v^T \frac{\partial g}{\partial x} = 0 \qquad \text{stationarity}$$

$$\frac{\partial L}{\partial v} = g(x) = 0 \qquad \text{feasibility}$$

replaced $\frac{\partial J}{\partial x} = 0 \Longrightarrow \frac{\partial L}{\partial x} = 0 \quad \frac{\partial L}{\partial v} = 0$

FROM ABOVE:

$$\min_x \ \frac{1}{2}(\tilde{y}_1 - H_1 x)^T W_1 (\tilde{y}_1 - H_1 x)$$

$$\text{s.t.} \quad \tilde{y}_2 = H_2 x \leftarrow$$

$$L(x, v) = \frac{1}{2}(\tilde{y}_1 - H_1 x)^T W_1 (\tilde{y}_1 - H_1 x) + v^T(\tilde{y}_2 - H_2 x)$$

$$\frac{\partial L}{\partial x} = -(\tilde{y}_1 - H_1 x)^T W_1 H_1 - v^T H_2 = 0$$

$$\frac{\partial L}{\partial v} = \tilde{y}_2 - H_2 x = 0 \leftarrow$$

solve for x & v

$$-\tilde{y}_1^T W_1 H_1 + x^T H_1^T W_1 H_1 = v^T H_2 \qquad \times (H_1^T W_1 H_1)^{-1} H_2^T$$

$$-\tilde{y}_1^T W_1 H_1 (H_1^T W_1 H_1)^{-1} H_2^T + \underbrace{x^T H_2^T}_{\tilde{y}_2^T} = v^T H_2 (H_1^T W_1 H_1)^{-1} H_2^T$$

note: H_1 tall H_2 fat.

$$v^T = \left(-\tilde{y}_1^T W_1 H_1 (H_1^T W_1 H_1)^{-1} H_2^T + \tilde{y}_2^T\right)\left(H_2 (H_1^T W_1 H_1)^{-1} H_2^T\right)^{-1}$$

$$x^T = \Big[\underbrace{\left(-\tilde{y}_1^T W_1 H_1 (H_1^T W_1 H_1)^{-1} H_2^T + \tilde{y}_2^T\right)\left(H_2 (H_1^T W_1 H_1)^{-1} H_2^T\right)^{-1} H_2}_{\text{addition from constraints.}} + \underbrace{\tilde{y}_1^T W_1 H_1\Big] (H_1^T W_1 H_1)^{-1}}_{\text{original LS solution}}$$

DERIVATIVE:

$$\frac{\partial f}{\partial x} : \Delta x \longmapsto \Delta f \qquad \Delta f = \frac{\partial f}{\partial x} \Delta x$$

Ex. $f(x) = c^T x \qquad \Delta f = c^T \Delta x \Rightarrow \frac{\partial f}{\partial x} = c^T$

$f(x) = Ax \qquad \Delta f = A \Delta x \Rightarrow \frac{\partial f}{\partial x} = A$

$f(x) = x^T Q x$

perturb ea. variable separately & add

$$\boxed{\Delta f = \Delta x^T Q x + x^T Q \Delta x}$$

often use trace

$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$

$\mathrm{Tr}(A) = \mathrm{Tr}(A^T)$

$$\Delta f = x^T Q^T \Delta x + x^T Q \Delta x \leftarrow$$

$$= x^T \underbrace{(Q^T + Q)}\, \Delta x$$

$$\frac{\partial f}{\partial x} = 2x^T Q \leftarrow Q \text{ sym.}$$