Least-Norm Solutions of Underdetermined Equations

Gabriel Hugh Elkaim

Least-Norm Solutions of Underdetermined Equations

- Least-norm solution of underdetermined equations
- Minimum norm solutions via QR factorization
- Derivation via Lagrange multipliers
- Relation to regularized least-squares
- General norm minimization for equality constraints

Underdetermined Linear Eq'ns (1.2)

$y = Ax$ each equation is a specification

$A \in \mathbb{R}^{m \times n}$ is <u>Fat</u> $(m < n)$

There are fewer equations than unknowns

x is underspecified, many choices for x give the same y.

Assume A is full rank (m)

for each $y \in \mathbb{R}^m$ there is an exact solution

Underdetermined Linear Eq'ns (2.2)

$$\{x \mid Ax = y\} = \{x_p + z \mid z \in N(A)\}$$

x_p – "particular solution" $Ax_p = y$.

In DESIGN this is GREAT NEWS

Solution has dim $N(A) = n - m$ "degrees of freedom"

Choose z to satisfy some OTHER specification

or do optimize something else.

Least-norm Solution (1.3)

$x_{ln} = \underbrace{A^T(AA^T)^{-1}}_{\text{right inverse of } A}\, y$ ← least norm solution

$x_{ls} = (A^TA)^{-1}A^T y$

$\left([A^T][A]\right)^{-1}$

$\left([A][A^T]\right)^{-1}$

(AA^T) invertible because A full rank (m)

LS: $y \approx Ax$ (approximate)

LN: $y = Ax$ (exactly)

x_{ln} is solution $y = Ax$ w/ $\|x\|^2$ minimized

Least-norm Solution (2.3)

$$\text{LN}\begin{cases}\min_x \|x\|^2 \\ \text{subj } Ax = y\end{cases}$$

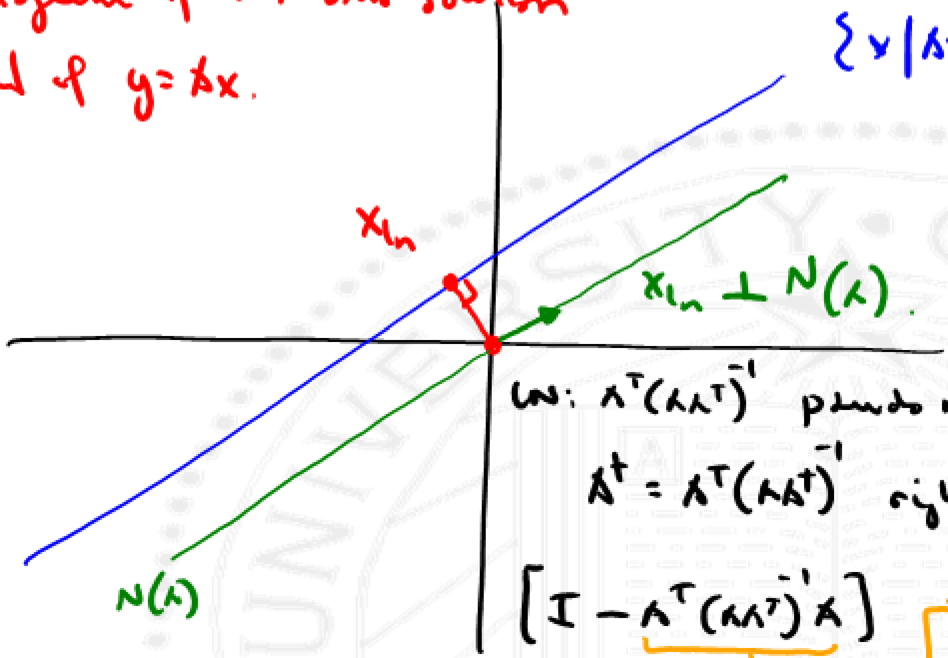
$$\text{LS}:\begin{cases}\min_x \|Ax-y\|^2\end{cases}$$

$$Ax = y \qquad A(x - x_{ln}) = 0$$

$$(x - x_{ln})^T x_{ln} = \underbrace{(x - x_{ln})^T A^T}_{0} (AA^T)^{-1} y$$

$$\boxed{x - x_{ln} \perp x_{ln}}$$

x_{ln} projection of (0) onto solution set of $y = Ax$.



Ls: $A^T(AA^T)^{-1}$ pseudo inverse of FAT A.

$A^{\dagger} = A^T(AA^T)^{-1}$ right inverse of A.

$\left[I - A^T(AA^T)^{-1}A \right] \rightarrow Pr(N(A))$

Least-norm Solution (3.3)

$$\|x\|^2 = \|x_{ln} + x - x_{ln}\|^2 = \|x_{ln}\|^2 + \overset{0}{\cancel{2x_{ln}^T(x - x_{ln})}} + \|x - x_{ln}\|^2$$

$$\geq \|x_{ln}\|^2$$

$$\|x_{ln}\|^2 + \|x - x_{ln}\|^2 \geq \|x_{ln}\|^2$$

x_{ln} has the smallest norm of all solutions to $y = Ax$.

Least-norm Solution via QR (1.3)

find QR of A^T — $A^T = QR$ (G-S on A^T)

$A^T = QR \qquad Q \in \mathbb{R}^{n \times m} \qquad Q^T Q = I_{m \times m}$

$R \in \mathbb{R}^{m \times m}$ upper triangular, invertible

$$x_{ln} = A^T (AA^T)^{-1} y = QR(R^T \underbrace{Q^T Q}_{I} R)^{-1} y = QR(R^T R)^{-1} = QRR^{-1}R^{-T}$$

$$x_{ln} = QR^{-T} \qquad \|x_{ln}\| = \|R^{-T} y\| \qquad R^{-T} \triangleq (R^T)^{-1}$$

Least-norm Solution via QR (2.3)

if A is FAT and full rank (m) ... plenty of solutions

why choose min $\|x\|^2$.

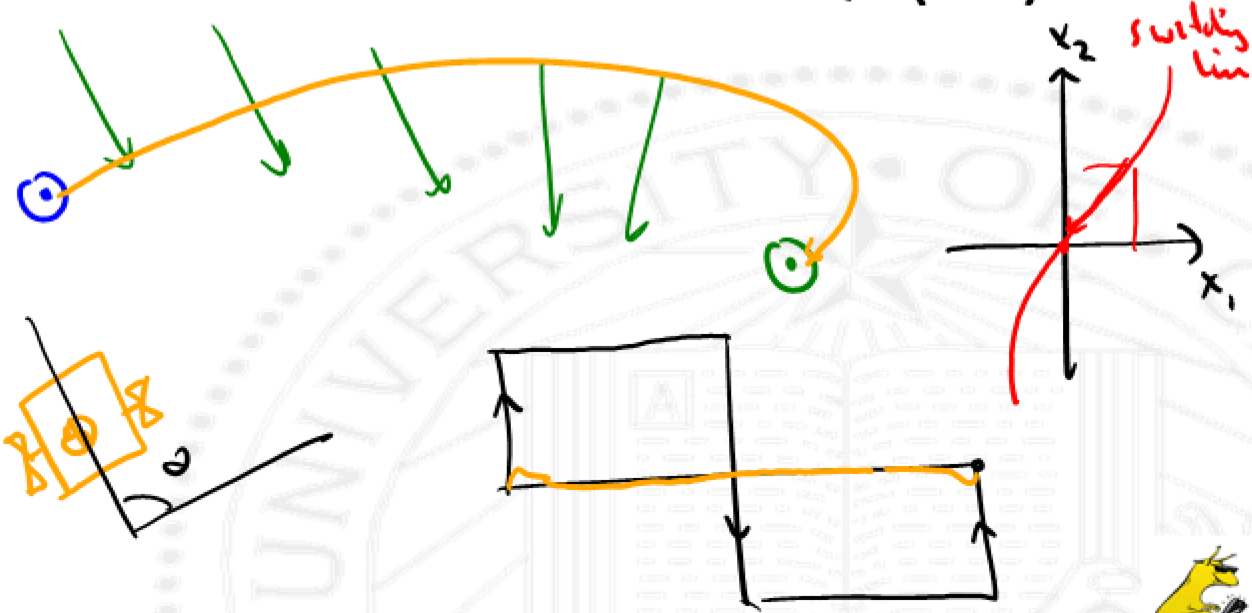
EASY TO CALCULATE

min $\Sigma |x_i|$ – 1 norm (FUEL BURN)

min max $|x_i|$ – ∞ norm (THRUSTER SIZE)

H_2 control (LQR, LQE, LQG)

Least-norm Solution via QR (3.3)



Derivation via Lagrange Multipliers

min $x^T x$

subj $Ax = y$

$$L(x,\lambda) = x^T x + \lambda(Ax - y)$$

optimality conditions are: $\frac{\partial L}{\partial x} = 0$ $\frac{\partial L}{\partial \lambda} = 0$.

$$\frac{\partial L}{\partial x} = 2x^T + \lambda A = 2x + A^T\lambda = 0 \longrightarrow x = -\frac{A^T\lambda}{2}$$

$$\frac{\partial L}{\partial \lambda} = Ax - y = 0 \longrightarrow A\left(-\frac{A^T\lambda}{2}\right) - y = 0 \longrightarrow -AA^T \cdot \frac{\lambda}{2} = y$$

$$\therefore \lambda = -2(AA^T)^{-1} y$$

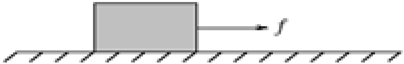
Derivation via Lagrange Multipliers

$x = -A^T \frac{\lambda}{2}$ $\qquad\qquad$ $\lambda = -2(AA^T)^{-1} y$

$\therefore \quad x = -A^T \frac{\lambda}{2} = \frac{-A^T(+2(AA^T)^{-1}}{2} y)$

$$\boxed{x_{ln} = A^T (AA^T)^{-1} y}$$

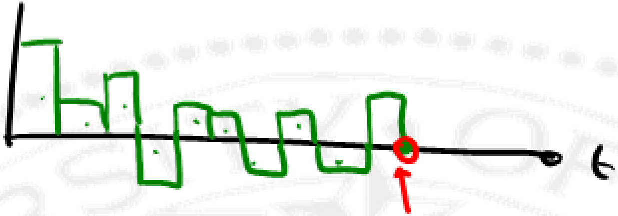
Lagrange Multiplier: Example (1.3)



y_1 position @ $t=10$

y_2 velocity @ $t=10$

Find a least norm mass transfer to $y_{des} = \begin{bmatrix}1\\0\end{bmatrix}$

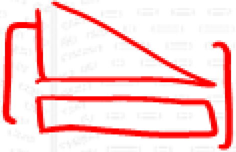


$y = Ax$

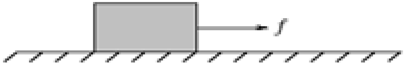
$A \in \mathbb{R}^{2\times 10}$

$x \in \mathbb{R}^{10}$

A is FAT

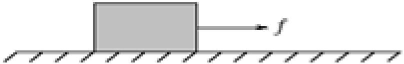


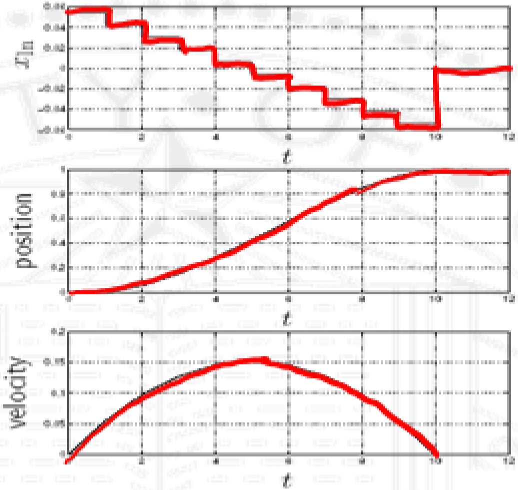
Lagrange Multiplier: Example (2.3)



$$x_{ln} = A^T(AA^T)^{-1}\begin{bmatrix}1\\0\end{bmatrix}$$

Lagrange Multiplier: Example (3.3)

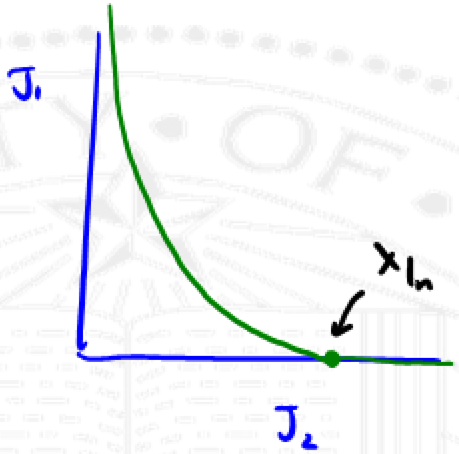




Relation to Regularized L-S (1.2)

$A \in \mathbb{R}^{m \times n}$ FAT, FULL RANK

$\min \; J_1 = \|Ax - y\|^2 \qquad J_2 = \|x\|^2$



LEAST NORM: $\min J_2$ w/ $J_1 = 0$.

$$\min (J_1 + \mu J_2) = \|Ax - y\|^2 + \mu \|x\|^2$$

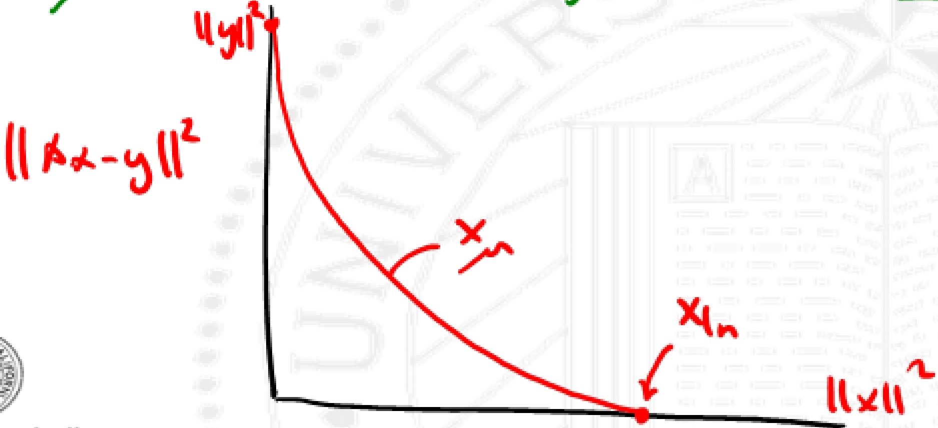
$$x_\mu = (A^T A + \mu I)^{-1} A^T y$$

$\boxed{x_\mu \longrightarrow x_{ln} \text{ as } \mu \to 0.}$

Relation to Regularized L-S (2.2)

as $\mu \to 0$

$(A^TA + \mu I)^{-1} A^T \longrightarrow A^T(AA^T)^{-1}$ for A is FAT full rank (m)



General Norm Minimization with Equality Constraints (1.3)

$$\min_x \|Ax-b\|$$

$$\text{subj to: } Cx = d$$

$$\min_x \tfrac{1}{2}\|Ax-b\|^2$$

$$\text{subj. } Cx = d$$

General Norm Minimization with Equality Constraints (2.3)

$$L(x,\lambda) = \frac{1}{2}\|Ax-b\|^2 + \lambda^T(Cx-d)$$

$$= \frac{1}{2}x^TA^TAx - b^TAx + \frac{1}{2}b^Tb + \lambda^TCx - \lambda^Td$$

$$\frac{\partial L}{\partial x} = 0 = A^TAx - A^Tb + C^T\lambda = 0 \quad \longrightarrow \text{soln for } \lambda.$$

$$\frac{\partial L}{\partial \lambda} = 0 = Cx - d = 0 \quad \longrightarrow \text{soln for } x\text{, substitute in } \lambda.$$

General Norm Minimization with Equality Constraints (3.3)

$$\left[\begin{array}{c|c} A^TA & C^T \\ \hline C & 0 \end{array}\right]\left[\begin{array}{c} x \\ \hline \lambda \end{array}\right] = \left[\begin{array}{c} A^Tb \\ \hline d \end{array}\right]$$

<u>KKT</u>

$$\left[\begin{array}{c} x \\ \hline \lambda \end{array}\right] = \left[\begin{array}{c|c} A^TA & C^T \\ \hline C & 0 \end{array}\right]^{-1}\left[\begin{array}{c} A^Tb \\ \hline d \end{array}\right]$$

if (A^TA) is invertible

$$x = (A^TA)^{-1}(A^Tb - C^T\lambda) \qquad Cx = d$$

$$C(A^TA)^{-1}(A^Tb - C^T\lambda) = d$$

$$\lambda = \left[C(A^TA)^{-1}C^T\right]^{-1}\left[C(A^TA)^{-1}A^Tb - d\right]$$

solve for x — UGLY

$$\min_x \|Ax - b\|$$

subj. $Cx = d$

LN: $A = I$, $b = 0$

$C = A$ $\quad d = y$.

$$\left[\begin{array}{c|c} I & A^T \\ \hline A & 0 \end{array}\right]\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

$$\boxed{x_{ln} = A^T (AA^T)^{-1} y.}$$

LS: $C = \{\}$

$d = \{\}$

$$\left[\begin{array}{c|c} A^T A & \phi \\ \hline \phi & \phi \end{array}\right]\begin{bmatrix} x \\ \hline \lambda \end{bmatrix} = \begin{bmatrix} A^T y \\ \hline \phi \end{bmatrix}$$

$$A^T A x = A^T y \longrightarrow \boxed{x_{ls} = (A^T A)^{-1} A^T y}$$

Questions?

this ends subject matter on MIDTERM

24 hour take home
5 problems

Open books, open notes, open video lectures (from any year)

NO LIVING HELP

NO PUBLISHED SOLUTIONS