Plan for the remainder of ESE 415:

- "Least-squares theory"

$$\begin{pmatrix} 1, 1 \\ 2, 2 \\ 3, 3 \\ \vdots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 12 \\ 18 \\ \vdots \end{pmatrix}$$

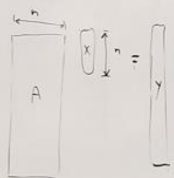
$Ax = y$ ← has no exact solution

Overdetermined linear system

$\min_{x \in \mathbb{R}^n} \|Ax - y\|^2$

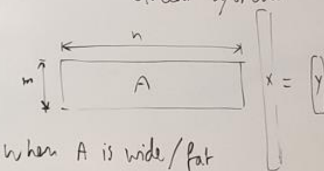
when A is tall

$m \gg n$



$Ax = y$ ← has infinitely many solutions

underdetermined linear linear system

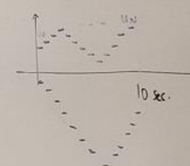


when A is wide/fat

$m \ll n$

$\min \|x\|$

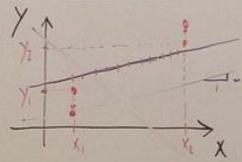
s.t. $Ax = y$



"minimum energy control solution"

- Convex Optimization with Constraints:
  - Lagrangian / Lagrange multipliers
  - Karush-Kuhn-Tucker condition (KKT)

Example for overdetermined case:



trivial case:

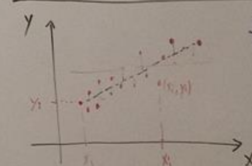
find straight line runs through (x_i, y_i), $i = 1, 2$

$f(x) = mx + c$

Solutions:

$$\left.\begin{matrix} f(x_1) = y_1 \\ f(x_2) = y_2 \end{matrix}\right\} \Leftrightarrow \begin{matrix} mx_1 + c = y_1 \\ mx_2 + c = y_2 \end{matrix} \Leftrightarrow \underbrace{\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} m \\ c \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}}_{y}$$

nontrivial (more interesting):



$f(x) = mx + c$

$$\underbrace{\begin{pmatrix} x_1, 1 \\ x_2, 1 \\ x_3, 1 \\ \vdots \\ x_N, 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} m \\ c \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix}}_{y}$$

$\min \|Ax - y\|^2$

$(1, 0, \cdots, 0)$

$A^T x \ldots$

$Ax = y$ is not solvable due to inconsistencies in the "data matrix" A

back-of-the-envelope derivation of least squares solution in this case (mnemonic device)

$$\underbrace{(A^T A)}_{(n \times n)} \underbrace{x}_{(n \times 1)} = \underbrace{A^T y}_{(n \times 1)}$$

unique solution to the NORMAL EQUATIONS

if $A^T A$ invertible ($A^T A \succ 0$):

$$\Rightarrow x^* = \underbrace{(A^T A)^{-1} A^T}_{=: A^\dagger} y$$

MATLAB: pinv(A) $=: A^\dagger$ PSEUDO-INVERSE (MOORE-PENROSE INVERSE)

Least squares Optimization as a subclass of Convex Optimization problems:

Def. A least-squares problem is an unconstrained optimization problem with an objective function

$$f(x) = \frac{1}{2}\|\underset{m\times n}{A}\,x - \underset{\in\mathbb{R}^m}{b}\|^2 = \frac{1}{2}(Ax-b)^T(Ax-b) = \dots = \underbrace{\frac{1}{2}x^TA^TA\,x - b^TAx + \frac{1}{2}b^Tb}_{\text{purely quadratic cost functional}}$$

$x \in \mathbb{R}^n$, 2-norm, constant

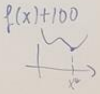
We implicitly assume that $A^TA \succ 0$.

$$\boxed{f(x) = \frac{1}{2}x^TQx - c^Tx + d}$$

general quadratic COST

min $f(x)$

min $f(x)+100$



Theorem: The solution $x^* \in \mathbb{R}^n$ of the least-squares problem satisfies

$$A^TAx^* = A^Tb \Rightarrow \text{If } A^TA \succ 0,\ x^* = (A^TA)^{-1}A^Tb.$$

Proof: Cost is convex, so consider $(\nabla f)(x) = \underbrace{A^TAx - A^Tb = A^T(Ax-b) = 0}_{A^TAx = A^Tb}$ solution to this is global minimizer.

Variants of Least-squares:
- weighted LS
- regularized LS: $f(x) = \frac{1}{2}\|Ax-b\|^2 + \lambda\, r(x)$ ← regularizing function,

regularization parameter

e.g.
- $r(x) = \|x\|^2$ (simplest example)
- $r(x) = \|x\|_1$ (Lasso) ← ℓ_1-norm

"Tikhonov regularizations"

Remark: Note that most formulations of LS lead to unconstrained quadratic programs (QP):

$$\text{minimize } \frac{1}{2}x^T\underset{\text{Symm., pos. def.}}{Q}x - c^Tx + d$$

Solution x^* to (QP) satisfies the NORMAL EQUATIONS

$$Qx^* - c = 0 \Rightarrow x^* = Q^{-1}c \quad (Q \succ 0).$$