EECS 275 Matrix Computation

Ming-Hsuan Yang

Electrical Engineering and Computer Science
University of California at Merced
Merced, CA 95344
http://faculty.ucmerced.edu/mhyang

Lecture 24

Overview

- Least-norm solutions of underdetermined equations
- General norm minimization with equality constraints

Underdetermined linear equations

- Consider $\mathbf{y} = A\mathbf{x}$ where $A \in \mathbb{R}^{m \times n}$ is fat $(m < n)$ i.e.,
  - there are more variables than equations
  - $\mathbf{x}$ is underdetermined, i.e., many choices $\mathbf{x}$ lead to the same $\mathbf{y}$
- Assume A is full rank (m), so each $\mathbf{y} \in \mathbb{R}^m$, there is a solution set of all solutions has form

$$\{\mathbf{x} | A\mathbf{x} = \mathbf{y}\} = \{\mathbf{x}_p + \mathbf{z} | \mathbf{z} \in \mathcal{N}(A)\}$$

  where $\mathbf{x}_p$ is any (particular) solution i.e., $A\mathbf{x}_p = \mathbf{y}$ and $\mathcal{N}(A)$ is the null space of A
- $\mathbf{z}$ characterizes available choices in solution
- A solution has $\dim(\mathcal{N}(A)) = n - m$ degrees of freedom
- Can choose $\mathbf{z}$ to satisfy other specifications or optimize among solutions

Least-norm solutions

- One particular solution is

$$\mathbf{x}_{ln} = A^\top (AA^\top)^{-1}\mathbf{y}$$

  ($AA^\top$ is invertible since A is full rank)

- $\mathbf{x}_{ln}$ is the solution $\mathbf{y} = A\mathbf{x}$ that minimizes $\|\mathbf{x}\|_2$, i.e., $\mathbf{x}_{ln}$ is solution of optimization problem

$$\begin{array}{ll} \text{min} & \|\mathbf{x}\|_2 \\ \text{subject to} & A\mathbf{x} = \mathbf{y} \end{array}$$

  with available $\mathbf{x} \in \mathbb{R}^n$

Least-norm solution (cont'd)

- One particular solution for $A\mathbf{x} = \mathbf{y}$ is

$$\mathbf{x}_{ln} = A^\top (AA^\top)^{-1}\mathbf{y}$$

- Suppose $A\mathbf{x} = \mathbf{y}$, so $A(\mathbf{x} - \mathbf{x}_{ln}) = 0$, and

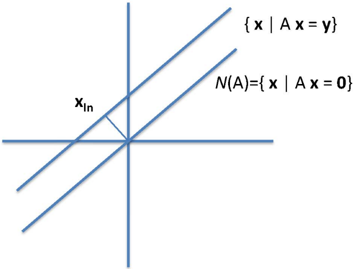
$$\begin{aligned} (\mathbf{x} - \mathbf{x}_{ln})^\top \mathbf{x}_{ln} &= (\mathbf{x} - \mathbf{x}_{ln})^\top A^\top (AA^\top)^{-1}\mathbf{y} \\ &= (A(\mathbf{x} - \mathbf{x}_{ln}))^\top (AA^\top)^{-1}\mathbf{y} \\ &= 0 \end{aligned}$$

i.e., $(\mathbf{x} - \mathbf{x}_{ln}) \perp \mathbf{x}_{ln}$, so

$$\|\mathbf{x}\|^2 = \|\mathbf{x}_{ln} + \mathbf{x} - \mathbf{x}_{ln}\|^2 = \|\mathbf{x}_{ln}\|^2 + \|\mathbf{x} - \mathbf{x}_{ln}\|^2 \geq \|\mathbf{x}_{ln}\|^2$$

i.e., $\mathbf{x}_{ln}$ has the smallest norm of any solution

Geometric interpretation



- Orthogonality condition: $\mathbf{x}_{ln} \perp \mathcal{N}(A)$
- Projection interpretation: $\mathbf{x}_{ln}$ is projection of $\mathbf{0}$ on solution set $\{\mathbf{x} | A\mathbf{x} = \mathbf{y}\}$
- $A^\dagger = A^\top (AA^\top)^{-1}$ is called the pseudoinverse of full rank, fat A
- $A^\top (AA^\top)^{-1}$ is a right inverse of A, i.e., $A \underbrace{A^\top (AA^\top)^{-1}}_{A^\dagger} = I$
- $I - A^\top (AA^\top)^{-1} A$ gives projection onto $\mathcal{N}(A)$

Overconstrained and underconstrained linear equations

Overconstrained	Underconstrained
$\min \; \|A\mathbf{x} - \mathbf{y}\|^2$	$\min \; \|\mathbf{x}\|$ subject to $A\mathbf{x} = \mathbf{y}$
$A\mathbf{x} = \mathbf{y}, \; m > n$	$A\mathbf{x} = \mathbf{y}, \; m < n$
$A^\dagger = (A^\top A)^{-1} A^\top$	$A^\dagger = A^\top (AA^\top)^{-1}$
$(A^\top A)^{-1} A^\top$ is a left inverse of A	$A^\top (AA^\top)^{-1}$ is a right inverse of A
$A(A^\top A)^{-1} A^\top$ gives projection onto $\mathcal{R}(A)$	$I - A^\top (AA^\top)^{-1} A$ gives projection onto $\mathcal{N}(A)$

Least-norm solution via QR factorization

- Find QR factorization of $A^\top$, i.e., $A^\top = QR$ with
  - $Q \in \mathbb{R}^{n \times m}$, $Q^\top Q = I_m$
  - $R \in \mathbb{R}^{m \times m}$, upper triangular, nonsingular
- $\mathbf{x}_{ln} = A^\top (AA^\top)^{-1} \mathbf{y} = QR^{-\top} \mathbf{y}$
- $\|\mathbf{x}_{ln}\| = \|R^{-\top} \mathbf{y}\|$

Derivation via Lagrange multipliers

- Least-norm solution solves optimization problem ($\|\mathbf{x}\|_2^2 = \mathbf{x}^\top \mathbf{x}$)

$$\begin{array}{ll} \min & \mathbf{x}^\top \mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{y} \end{array}$$

- Introduce Lagrange multiplers: $L(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{x}^\top \mathbf{x} + \boldsymbol{\lambda}^\top (A\mathbf{x} - \mathbf{y})$
- Optimality conditions are

$$\nabla_{\mathbf{x}} L = 2\mathbf{x} + A^\top \boldsymbol{\lambda} = 0, \quad \nabla_{\boldsymbol{\lambda}} L = A\mathbf{x} - \mathbf{y} = 0$$

- From the first condition, $\mathbf{x} = -A^\top \boldsymbol{\lambda}/2$
- Substitute into the second condition, $\boldsymbol{\lambda} = -2(AA^\top)^{-1}\mathbf{y}$
- Hence $\mathbf{x} = A^\top (AA^\top)^{-1}\mathbf{y}$

Relation to regularized least-squares

- Suppose $A \in \mathbb{R}^{m \times n}$ is fat, full rank
- Define $J_1 = \|A\mathbf{x} - \mathbf{y}\|^2$, $J_2 = \|\mathbf{x}\|^2$
- Least-norm solution minimizes J_2 with $J_1 = 0$
- Minimizer of weighted-sum objective $J_1 + \mu J_2 = \|A\mathbf{x} - \mathbf{y}\|^2 + \mu\|\mathbf{x}\|^2$ is

$$\mathbf{x}_\mu = (A^\top A + \mu I)^{-1} A^\top \mathbf{y}$$

- Fact: $\mathbf{x}_\mu \to \mathbf{x}_{ln}$ as $\mu \to 0$, i.e., regularized solution converges to least-norm solution as $\mu \to 0$
- In matrix form, as $\mu \to 0$

$$(A^\top A + \mu I)^{-1} A^\top \to A^\top (AA^\top)^{-1}$$

for full rank, fat A

General norm minimization with equality constraints

- Consider the problem

$$\begin{array}{ll} \min & \|A\mathbf{x} - \mathbf{b}\| \\ \text{subject to} & C\mathbf{x} = \mathbf{d} \end{array}$$

  with variable $\mathbf{x}$
- Includes least-squares and least-norm problems as special cases
- Equivalent to

$$\begin{array}{ll} \min & \frac{1}{2}\|A\mathbf{x} - \mathbf{b}\|^2 \\ \text{subject to} & C\mathbf{x} = \mathbf{d} \end{array}$$

- Lagrangian is

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}) &= \frac{1}{2}\|A\mathbf{x} - \mathbf{b}\|^2 + \boldsymbol{\lambda}^\top (C\mathbf{x} - \mathbf{d}) \\ &= \frac{1}{2}\mathbf{x}^\top A^\top A\mathbf{x} - \mathbf{b}^\top A\mathbf{x} + \frac{1}{2}\mathbf{b}^\top \mathbf{b} + \boldsymbol{\lambda}^\top C\mathbf{x} - \boldsymbol{\lambda}^\top \mathbf{d} \end{aligned}$$

Norm minimization with equality constraints (cont'd)

- Lagrangian is

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}) &= \tfrac{1}{2}\|A\mathbf{x} - \mathbf{b}\|^2 + \boldsymbol{\lambda}^\top (C\mathbf{x} - \mathbf{d}) \\ &= \tfrac{1}{2}\mathbf{x}^\top A^\top A\mathbf{x} - \mathbf{b}^\top A\mathbf{x} + \tfrac{1}{2}\mathbf{b}^\top \mathbf{b} + \boldsymbol{\lambda}^\top C\mathbf{x} - \boldsymbol{\lambda}^\top \mathbf{d} \end{aligned}$$

- Optimality conditions are

$$\nabla_{\mathbf{x}} L = A^\top A\mathbf{x} - A^\top \mathbf{b} + C^\top \boldsymbol{\lambda} = 0, \quad \nabla_{\boldsymbol{\lambda}} L = C\mathbf{x} - \mathbf{d} = 0$$

- Put in matrix form

$$\begin{bmatrix} A^\top A & C^\top \\ C & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} A^\top \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

- If the block matrix is invertible, we have

$$\begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} A^\top A & C^\top \\ C & 0 \end{bmatrix}^{-1} \begin{bmatrix} A^\top \mathbf{b} \\ \mathbf{d} \end{bmatrix}$$

Norm minimization with equality constraints (cont'd)

- If $A^\top A$ is invertible, we can derive a more explicit formula for $\mathbf{x}$
- From the first block equation, we have

$$\mathbf{x} = (A^\top A)^{-1}(A^\top \mathbf{b} - C^\top \boldsymbol{\lambda})$$

- Substitute into $C\mathbf{x} = \mathbf{d}$

$$C(A^\top A)^{-1}(A^\top \mathbf{b} - C^\top \boldsymbol{\lambda}) = \mathbf{d}$$

so

$$\boldsymbol{\lambda} = (C(A^\top A)^{-1}C^\top)^{-1}(C(A^\top A)^{-1}A^\top \mathbf{b} - \mathbf{d})$$

- Recover $\mathbf{x}$ from equation above

$$\mathbf{x} = (A^\top A)^{-1}(A^\top \mathbf{b} - C^\top (C(A^\top A)^{-1}C^\top)^{-1}(C(A^\top A)^{-1}A^\top \mathbf{b} - \mathbf{d}))$$