Video Lecture G7: Least-Squares Approximation to Solving Linear Systems

Tom Roby

Outline & Objectives

- Utilize the machinery of orthogonal projections to find *least-squares solutions* to inconsistent linear systems $A\vec{x} = \vec{b}$, where $\vec{b} \notin \mathrm{Col}\, A$. Such solutions are the best approximations to actual solutions, and they are frequently needed in applications.

- Investigate the special cases where the columns of A are orthogonal, or we have a QR factorization of A.

Least-Squares Solutions to $A\vec{x} = \vec{b}$.

Definition (least-squares solutions)

For $A \in \mathbb{R}^{m\times n}$ and $\vec{b} \in \mathbb{R}^m$, a least-squares solution $\widehat{x} \in \mathbb{R}^n$ to $A\vec{x} = \vec{b}$ satisfies $\|\vec{b} - A\widehat{x}\| \le \|\vec{b} - A\vec{x}\|\ \forall\, \vec{x} \in \mathbb{R}^n$.

Theorem (Normal equations)

$\widehat{x}$ is a least-squares soln to $A\vec{x} = \vec{b} \iff A^T A\widehat{x} = A^T\vec{b}$ (the normal equations for $A\vec{x} = \vec{b}$), which always has at least one soln.

Proof: Let $A = [\vec{a}_1 \ \cdots \ \vec{a}_n]$ and $\widehat{b} = \mathrm{Proj}_{\mathrm{Col}\,A}\,\vec{b}$. Then $\exists\, \widehat{x}$ such that $A\widehat{x} = \widehat{b}$. By Orthogonal Proj Thm, $\vec{b} - \widehat{b} = \vec{b} - A\widehat{x}$ is in $(\mathrm{Col}\,A)^\perp$, equivalently, $\vec{a}_j \cdot (\vec{b} - A\widehat{x}) = 0$ for $j \in [n]$. Rewrite this in matrix form to get $A^T(\vec{b} - A\widehat{x}) = \vec{0} \iff A^T A\widehat{x} = A^T\vec{b}$. ■

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 2 & 3 \end{bmatrix}, \ \vec{b} = \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix}.$$

Theorem (Uniqueness of Least-Squares Solns)

For $A \in \mathbb{R}^{m\times n}$ TFAE: (a) $A\vec{x} = \vec{b}$ has a unique LS Soln; (b) $A^T A$ is invertible; (c) Columns of A are lin indep.

Proof: Normal Equation Thm shows (a) $\iff$ (b), rest exercise. ■

What if columns of A are orthogonal? Then $\widehat{b}$ and $\widehat{x}$ are easy.

$$A = \begin{bmatrix} 1 & -3 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 4 \\ 5 \\ 0 \\ -1 \end{bmatrix}.$$

Theorem (Using QR-factorization for least-squares)

Let $A \in \mathbb{R}^{m\times n}$ have rank n (lin indep cols) and QR factorization $A = QR$. Then for any $\vec{b} \in \mathbb{R}^m$, the unique LS soln to $A\vec{x} = \vec{b}$ is given by $\widehat{x} = R^{-1}Q^T\vec{b}$.

Proof: $A\widehat{x} = (QR)(R^{-1}Q^T\vec{b}) = QQ^T\vec{b} = \mathrm{Proj}_{\mathrm{Col}\,A}\,\vec{b} = \widehat{b}$. ■