Math 304

Linear Algebra

Harold P. Boas

boas@tamu.edu

June 20, 2006

Highlights

From last time:

- orthogonal subspaces

Today:

- least squares problems

Solving unsolvable problems

Recall that the linear system $A\mathbf{x} = \mathbf{b}$ is solvable (that is, consistent) if and only if the vector $\mathbf{b}$ belongs to the column space of the matrix A (that is, the range of A).

If the system is unsolvable (that is, inconsistent), we might ask for a vector $\mathbf{x}$ that *minimizes* the length of the difference vector $A\mathbf{x} - \mathbf{b}$. Such an $\mathbf{x}$ is a *least squares solution*.

Reinterpretation: the vector $A\mathbf{x} - \mathbf{b}$ should be orthogonal to $R(A)$.
We know from last time that $R(A)^\perp = N(A^T)$, so we want $A^T(A\mathbf{x} - \mathbf{b}) = \mathbf{0}$.
Thus we have a reformulation of the least squares problem:

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Example

Find the line of the form $y = a + bx$ that best fits the data:

x	0	1	2
y	1	2	5

[answer: $y = \frac{2}{3} + 2x$]

Solution. We seek a least-squares solution of the system:

$$\begin{aligned} a + b \times 0 &= 1 \\ a + b \times 1 &= 2 \\ a + b \times 2 &= 5 \end{aligned} \quad \text{or} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}.$$

Multiplying by the transpose $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ gives the system:

$$\begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix}.$$

Solve either by row reduction or by multiplying by $\begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}^{-1}$.

Remarks on $A^T A\mathbf{x} = A^T\mathbf{b}$

- $A^T A$ is always a *symmetric* square matrix.
- $A^T A\mathbf{x} = A^T\mathbf{b}$ is always a *consistent* system.
- There is a *unique* least squares solution if and only if $A^T A$ is an invertible matrix.
- There is a *unique* least squares solution if and only if the columns of A are linearly independent (that is, A is an $m \times n$ matrix of rank n).