MAT225 Section Summary: 6.5

Least-Squares Problems

Summary

Okay! This is it: the section with the formula for the solution of the least-squares problem, which is known as the linear regression problem in statistics. This is how we find a nice fit to linear (and specialized types of non-linear) models. What an amazingly powerful tool this is, and it's based on some simple linear algebra....

least-squares solution: If $A_{m \times n}$ and $\mathbf{b}$ is in $\mathbb{R}^m$, then a least-squares solution of $A\mathbf{x} = \mathbf{b}$ is $\hat{\mathbf{x}}$ in $\mathbb{R}^n$such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.

Q: Take a look at that equation above, and tell me where the name "least-squares" comes from....

Now, consider the projection of $\mathbf{b}$ onto the Col A,

$$\hat{\mathbf{b}} = \text{proj}_{\text{Col } A}\mathbf{b}$$

and let $\hat{\mathbf{x}}$ be defined as the solution of

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}}.$$

Q: How do we know that there **is** such a solution?

We know that $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to Col A, so

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}.$$

$$\underbrace{\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}}_{A} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

from which we arrive at

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

Hence $\hat{\mathbf{x}}$ is a solution of the equation

$$\boxed{A^TA\mathbf{x} = A^T\mathbf{b}}$$

$$A^TA_{2\times 2}\begin{bmatrix} m \\ b \end{bmatrix} = A^Tb_{2\times 1}$$

Solve for

$$\hat{x} = \begin{bmatrix} m \\ b \end{bmatrix}$$

(the so-called **normal equations**). There may be many (in infinite number!) of solutions of the normal equations.

Theorem 13: The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions of the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$.

However, if $A^T A$ is invertible, then the solution is unique:

Theorem 14: The matrix $A^T A$ is invertible $\iff$ the columns of A are linearly independent. In this case, the equation $A\mathbf{x} = \mathbf{b}$ has only one least-squares solution $\hat{\mathbf{x}}$, and it is

$$\boxed{\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}.}$$

Problems:

1. #2, p. 416

$$\begin{bmatrix}2&1\\-2&0\\2&3\end{bmatrix}\begin{bmatrix}m\\b\end{bmatrix}=\begin{bmatrix}-5\\8\\1\end{bmatrix}$$

$$A \quad \bar{x} = \bar{b}$$

$$A^T A = \begin{bmatrix}2&-2&2\\1&0&3\end{bmatrix}\begin{bmatrix}2&1\\-2&0\\2&3\end{bmatrix} = \begin{bmatrix}12&8\\8&10\end{bmatrix} \quad \text{(symmetric)}$$

$$A^T b = \begin{bmatrix}2&-2&2\\1&0&3\end{bmatrix}\begin{bmatrix}-5\\8\\1\end{bmatrix} = \begin{bmatrix}-24\\-2\end{bmatrix}$$

$$\begin{bmatrix}12&8\\8&10\end{bmatrix}\begin{bmatrix}\hat{m}\\\hat{b}\end{bmatrix}=\begin{bmatrix}-24\\-2\end{bmatrix}$$

$$\Rightarrow \begin{bmatrix}\hat{m}\\\hat{b}\end{bmatrix}$$

$$\hat{x} = \begin{bmatrix}-4\\3\end{bmatrix} \qquad \begin{bmatrix}12&8\\8&10\end{bmatrix}\begin{bmatrix}-4\\3\end{bmatrix}=\begin{bmatrix}-24\\-2\end{bmatrix} \checkmark$$

$$\begin{bmatrix}2&1\\-2&0\\2&3\end{bmatrix}\begin{bmatrix}-4\\3\end{bmatrix}=\begin{bmatrix}-5\\8\\1\end{bmatrix} \checkmark!$$

A miracle occurred: $\bar{b}$ was in the Col(A)

2. #5, p. 416

$$A = \begin{bmatrix}1&1&0\\1&1&0\\1&0&1\\1&0&1\end{bmatrix}$$

$$\bar{b} = \begin{bmatrix}1\\3\\8\\2\end{bmatrix}$$

linear dependence in the columns of $A \Rightarrow$ an infinite number of least squares solutions

3. #13, p. 416

$$A = \begin{bmatrix} 3 & 4 \\ -2 & 1 \\ 3 & 4 \end{bmatrix} \quad \bar{b} = \begin{bmatrix} 11 \\ -9 \\ 5 \end{bmatrix} \quad \bar{u} = \begin{bmatrix} 5 \\ -1 \end{bmatrix} \quad \bar{v} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Compute $A\bar{v}$ + $A\bar{u}$

$$A\bar{u} = \begin{bmatrix} 11 \\ -11 \\ 11 \end{bmatrix} \qquad A\bar{v} = \begin{bmatrix} 7 \\ -12 \\ 7 \end{bmatrix}$$

Compute (sum of squares)

$$\|A\bar{u} - \bar{b}\|^2 = 40$$

+ $\|A\bar{v} - \bar{b}\|^2 = 29$ ← less squares!

$A\bar{v}$ beats $A\bar{u}$ at approximating $\bar{b}$. & compare. $\bar{u}$ can't possibly be the least squares soln.

4. #23, p. 417

5. #24, p. 417

$A\bar{x} = \bar{b}$; least squares soln when columns of A are orthonormal.

$$\underbrace{A^T A}_{I} \hat{x} = A^T \bar{b}$$

I, so $\boxed{\hat{x} = A^T \bar{b}}$

6. #25, p. 417