MAT225 Section Summary: 6.3

Orthogonal Projections
Summary

This section formalizes one of the things that I've been emphasizing all along about projections, orthogonal complements, etc., to whit: we can't solve the equation $A\mathbf{x} = \mathbf{b}$, so we try to solve the next best thing: we solve $A\mathbf{x} = \hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is the projection of $\mathbf{b}$ onto the column space of A.

Theorem 8: The Orthogonal Decomposition Theorem Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$. In fact, if $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \ldots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p$$

and then $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

orthogonal projection of y onto W: The vector $\hat{\mathbf{y}}$ is called the orthogonal projection of y onto W, written $\text{proj}_W \mathbf{y}$.

Properties of orthogonal projections:

1. If $\mathbf{y}$ is in $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$, then $\text{proj}_W \mathbf{y} = \mathbf{y}$. $(\vec{z} = \vec{0})$
2. The orthogonal projection of y onto W is the best approximation to $\mathbf{y}$ by elements of W.

Theorem 9: The Best Approximation Theorem Let W be a subspace of $\mathbb{R}^n$, $\mathbf{y}$ any vector in $\mathbb{R}^n$, and $\hat{\mathbf{y}}$ the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

Theorem 10: If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \ldots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p$$

If $U = [\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \ldots \ \ \mathbf{u}_p]$, then

$$\text{proj}_W \mathbf{y} = UU^T\mathbf{y}$$

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

Now, as an example, I want to consider Taylor series expansions for function with three derivatives at a point a (that might define our space: you should check that this is indeed a vector space, by checking that it's a subspace of the space of thrice differentiable functions). The Taylor series expansion for the function f is

$$C(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2} + f'''(a)\frac{(x-a)^3}{6}$$

This is a vector in the space $\mathbb{P}_3$. What we're doing is projecting the vector f (which is otherwise unspecified) onto $\mathbf{P}_3$, in a way that minimizes the distance between the vectors

$$\begin{bmatrix} C(a) \\ C'(a) \\ C''(a) \\ C^{(3)}(a) \end{bmatrix} \text{ and } \begin{bmatrix} f(a) \\ f'(a) \\ f''(a) \\ f^{(3)}(a) \end{bmatrix}$$

(in fact, the difference between these vectors is zero!).

Now with functions you have to be a little careful, because it's a little tricky to define just what is meant by an inner-product. We're not going to get into that...!

\# 12 $\quad \bar{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix} \quad \bar{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix} \quad \bar{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$

Note: $\bar{v}_1 \perp \bar{v}_2 \quad (\bar{v}_1 \cdot \bar{v}_2 = 0)$

$$\hat{y} = \frac{\bar{y}\cdot\bar{v}_1}{\|\bar{v}_1\|^2}\bar{v}_1 + \frac{\bar{y}\cdot\bar{v}_2}{\|\bar{v}_2\|^2}\bar{v}_2 = \begin{bmatrix} 3 \\ -6 \\ -3 \\ 6 \end{bmatrix} + \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ -3 \\ 9 \end{bmatrix}$$