Section 6.3: Orthogonal Projections

April 18, 2008

Abstract

This section formalizes one of the things that I've been emphasizing all along about projections, orthogonal complements, etc., to whit: when we can't solve the equation $A\mathbf{x} = \mathbf{b}$ exactly, we 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

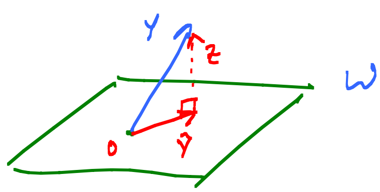
$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

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 $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. z is going to represent error in the least squares problem

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



Example: #1, p. 400

(Used Octave example)

Properties of orthogonal projections:

(a) If $\mathbf{y}$ is in $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$, then $\text{proj}_W\mathbf{y} = \mathbf{y}$.

(b) 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

$$\underbrace{\|\mathbf{y} - \hat{\mathbf{y}}\|}_{\|\underline{z}\|} \le \|\mathbf{y} - \mathbf{v}\|$$

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

Example: Revisit #1, p. 400

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

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

In the Octave example

$\underline{x}' \cdot U$

$U = [u_1\ u_2\ u_3\ u_4]$

(Alternatively get them by

$U' \cdot \underline{x}$

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

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

$$\text{proj}_W \underline{x} = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \underline{u}_3 \end{bmatrix} \begin{bmatrix} \underline{u}_1^T \\ \underline{u}_2^T \\ \underline{u}_3^T \end{bmatrix} \begin{bmatrix} 10 \\ -4 \\ 2 \\ 0 \end{bmatrix}$$

$\text{span}\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$

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

Example: Revisit #1, p. 400

Now for a completely different example: I want to consider Taylor series expansions for function with three derivatives at a point a (that property defines our vector 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 about a 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 P_3. What we're doing is projecting the vector f (which is otherwise unspecified) onto P_3, in a way that minimizes the distance between the vectors $p(x) \in P_3$ and $f(x)$. I'm asserting that $\|\mathbf{f(x)} - \mathbf{C(x)}\|$ is minimal among elements of P_3.

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 now...!