MATH 54 Lecture Notes 12

GSI Carter

July 17, 2007

1 Least Squares

Consider the vector space V of arbitrary functions from the set $\{-1, 0, 1\}$ to $\mathbb{R}$. This vector space has a basis $B = \{f_1, f_2, f_3\}$, where

$$f_1(-1) = f_2(0) = f_3(1) = 1$$

and

$$f_1(0) = f_1(1) = f_2(-1) = f_2(1) = f_3(-1) = f_3(0) = 0.$$

In particular, let f be an arbitrary element of V, where $f(-1) = a$, $f(0) = b$, and $f(1) = c$. Then

$$f = af_1 + bf_2 + cf_3.$$

Therefore

$$[f]_B = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

By using coordinate vectors relative to B, we can think of V as being the same as $\mathbb{R}^3$. Then we can take the standard dot product in $\mathbb{R}^3$ to get an inner product on V.

Now suppose we have the data points $(-1, 1)$, $(0, 1)$, and $(1, 2)$. This data is a function from $\{-1, 0, 1\}$ to $\mathbb{R}$. In other words, the data is an element of V, $f_1 + f_2 + 2f_3$. Then the coordinate vector relative to B of the data is $(1, 1, 2)^T$. We wish to approximate this data with a line, or an element of $\text{Span}\{1, x\}$. In V,

$$1 = f_1 + f_2 + f_3$$

and

$$x = f_3 - f_1.$$

Therefore we are trying to approximate $(1, 1, 2)^T$ by a point on some two-dimensional subspace of $\mathbb{R}^3$. This will be accomplished by choosing the perpendicular projection of the point onto this plane. If $\mathbf{b} = (1, 1, 2)^T$ and

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix},$$

then we are looking for $\mathbf{x} \in \mathbb{R}^2$ such that $A\mathbf{x}$ is the perpendicular projection of $\mathbf{b}$ onto $CS(A)$, or such that $A\mathbf{x} - \mathbf{b}$ is perpendicular to $CS(A)$. To do this, we solve the equation

$$A^T A\mathbf{x} = A^T\mathbf{b} \tag{1}$$

for $\mathbf{x}$.

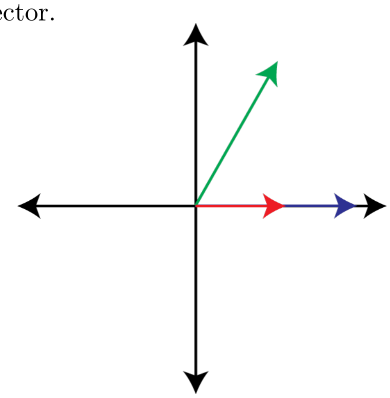
To see why this works, we use the fact that $(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A^T\mathbf{v})$, provided that all the entries of A are real. Then the following are equivalent:

- $A\mathbf{x} - \mathbf{b} \perp CS(A)$.
- $\langle A\mathbf{x} - \mathbf{b}, A\mathbf{y}\rangle$ for all $\mathbf{y}$.
- $\langle A^T(A\mathbf{x} - \mathbf{b}), \mathbf{y}\rangle$ for all $\mathbf{y}$.
- $A^T(A\mathbf{x} - \mathbf{b}) = 0$.

The equation (1) always has a solution since $CS(A^T A) \subseteq CS(A^T)$ (in general) and $rk(A^T A) = rk(A^T)$ (theorem in the book, p. 258), so that $CS(A^T A) = CS(A^T)$. Then since $A^T\mathbf{b} \in CS(A^T)$, there must be a solution.

2 Gram-Schmidt

Recall that in an inner product space, we define two vectors $\mathbf{u}$ and $\mathbf{v}$ to be perpendicular if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. In order to perform projections onto a multi-dimensional subspace, we will first need a basis for this subspace consisting of vectors which are all perpendicular to each other. We call such a basis *orthogonal*. We will use Gram-Schmidt to turn a basis into an orthogonal basis for the same vector space (that is, the span will be preserved, and the vectors will still be linearly independent, but now they will be perpendicular to each other). The idea behind Gram-Schmidt is to take each successive vector and subtract off the components which are in the same direction as each previous vector.



Suppose the blue vector is $\mathbf{v}_1$ and the green vector is $\mathbf{v}_2$. Then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^2$ to which we can apply Gram-Schmidt. Then the red vector is $\text{proj}_{\mathbf{v}_1}\mathbf{v}_2$, and $\mathbf{v}_2 - \text{proj}_{\mathbf{v}_1}\mathbf{v}_2$ is perpendicular to $\mathbf{v}_1$.

Given linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, the vectors

$$\begin{aligned}
\mathbf{p}_1 &= \mathbf{v}_1 \\
\mathbf{p}_2 &= \mathbf{v}_2 - \operatorname{proj}_{\mathbf{v}_1} \mathbf{v}_2 \\
\mathbf{p}_3 &= \mathbf{v}_3 - \operatorname{proj}_{\mathbf{v}_1} \mathbf{v}_3 - \operatorname{proj}_{\mathbf{v}_2} \mathbf{v}_3 \\
&\ \vdots \\
\mathbf{p}_n &= \mathbf{v}_n - \sum_{i=1}^{n-1} \operatorname{proj}_{\mathbf{v}_i} \mathbf{v}_n
\end{aligned}$$

are an orthogonal basis for $\operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

3 Multi-Dimensional Projections

Given an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for a subspace V of an inner product space W, then vectors in W can be orthogonally projected onto V by the following formula:

$$\operatorname{proj}_V \mathbf{w} = \operatorname{proj}_{\mathbf{v}_1} \mathbf{w} + \operatorname{proj}_{\mathbf{v}_2} \mathbf{w} + \cdots + \operatorname{proj}_{\mathbf{v}_n} \mathbf{w}.$$

We can use this, for example, to project vectors onto the plane $x + y + z = 0$.