Projection and Best Fit

Using matrix operation to accomplish projections and compute best fits

Projection and Best Fit

We will now put together all of our work on matrix operations, eventually to return to our motivating question from the beginning of class: *Given a collection of data, find a polynomial that fits that data* now with the added wrinkle of *as close as possible.* For example, given a set of height and weight pairs for 100 individuals, find the linear function that is closest to this data.

1. We begin with projection. Recall that the dot product of two vectors gives the cosine of the angle between those vectors. In matrix terms, given two column vectors we can write this as:

$$v \cdot w = w^T v = |w||v| \cos(\theta)$$

where the w^T is the transpose of w, found by changing rows to columns and vice versa. For vectors v and u in $\mathbb{R}^2$, and u with unit length ($u^T u = 1$), show that the length of the projection of v onto the line spanned by u is given by

$$v \cdot u = u^T v$$

2. Then show that $v = w+x$ where w is in the direction of u and x is perpendicular to u (which you will recall means $u^T x = 0$). The vector w is then the *orthogonal projection* of v onto the line spanned by u. For our class we will only discuss orthogonal projection, so we will usually leave off the *orthogonal*.

3. Finally note that if the vector we are projecting onto is not unit length our formula changes slightly: The project of v onto w is given by

$$\frac{w^T v}{w^T w} w$$

4. So a natural definition of projection in higher dimensions is then: Given a vector space V, and a subspace W with basis $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ the projection of $v \in V$ onto W is given by the vector

$$A \left(A^T A\right)^{-1} A^T v$$

where A is the $n \times m$ matrix formed by the columns $\alpha_1, \alpha_2, \ldots, \alpha_m$. Equivalently the projection is written as

$$\text{proj}_W(v) = c_1\alpha_1 + c_2\alpha_2 + \ldots c_m\alpha_m$$

where the c_j are given by

$$\left(A^T A\right)^{-1} A^T v$$

Show that if $m = n$ and $W = V$ then this formula gives the correct answer.

5. Given the subspace

$$M = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} | x + y + z = 0 \right\}$$

find the projection of $v = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ onto M.

6. A random smaple of 10 individuals gave the following heights and weights:

Height (inches)	Weight (pounds)
66	113
72	136
69	153
68	142
68	144
69	123
70	141
70	136
68	112
67	121

Table 1: Heights and weights of 10 individuals.

Set up the linear problem to determine the line describing the relationship between height and weight for this data set. Does this system have (a) a unique solution; (b) infinitely many solutions; or (c) no solutions?

7. Show that the linear system

$$Ax = b$$

has solutions only if b is in range of A (i.e. the span of the columns of A).

8. So the problem with our height-weight problem is that the b is not in the range. However we now know that we can project it to the range of A if the columns of A are independent (i.e. form a basis for a subspace). Then we have

$$x = \left(A^T A\right)^{-1} A^T b$$

Find the best fit linear equation for the height-weight data.