Projections

With the inner product $<x, y>$, we have angles ($<x, y> = \|x\|_2 \|y\|_2 \cos(\theta)$), and can speak of orthogonality: $x \perp y \iff <x, y> = 0$. Here we will consider the standard inner product for $\mathbb{R}^n$: $<x, y> \equiv x^T y$, but more general inner products can be very useful in many applications and algorithm development.

If S is a subspace of $\mathbb{R}^n$ (write $S \leq \mathbb{R}^n$), we say $x \perp S$ if x is orthogonal to every element of S. The subspace $S^\perp \leq \mathbb{R}^n$ is called the *orthogonal complement* of S

$$S^\perp \equiv \{x \in \mathbb{R}^n : x \perp S\},$$

and $\mathbb{R}^n = S \oplus S^\perp$ is a direct sum decomposition of $\mathbb{R}^n$ into complementary subspaces in such a way that each $x \in \mathbb{R}^n$ has the unique factorization $x = u + v$, with $u \in S$ and $v \in S^\perp$. In this setting, u is the *orthogonal projection* of x onto S.

In this note, we will be looking at a transformation P, which satisfies

$$\forall x \in \mathbb{R}^n,\ Px = u,\ \text{the orthogonal projection of } x \text{ onto } S.$$

If $x \in S$, then we should have $Px = x$ (right?), so P should satisfy $P^2 = P$, with Range$(P) = S$. The requirement that $Px = u \perp v$, with $u \in S$ and $v \in S^\perp$, combined with $x^T y = y^T x$ forces P to be self-adjoint (in matrix language over $\mathbb{R}$, $P^T = P$). Any linear transformation P which satisfies

1. Range$(P) = S$,
2. $P^2 = P$, and
3. $P^T = P$

is called an *orthogonal projector* onto S. If Q is another orthogonal projector onto S, then $Qx = Q(u + v) = u = Px,\ \forall x \in \mathbb{R}^n$, and hence $Q = P$ and we see that the orthogonal projector onto a subspace is unique. If we have a basis, then we should expect to be able to find a matrix representation for P; call it P. Let X have linearly independent columns spanning S. Now (scratch paper handy?)
$PX = X \Rightarrow X^T PX = X^T X$, Range$(P)$=Span$(X) \Rightarrow P = XM$ for some M, $P = P^T$ $\Rightarrow XM = M^T X^T$, $P = P^2 \Rightarrow P = XMM^T X^T$, giving $MM^T = (X^T X)^{-1}$ so that

$$P = X(X^T X)^{-1} X^T.$$

$X^T X$ is nonsingular, so the derivation and formula are perfectly reasonable, but if the columns of V form an *orthonormal basis* for S, then $V^T V = I$, in which case

$$P = VV^T.$$

Now from $x = Px \oplus v$, we see that $v = (I - P)x$, and (check the properties) the orthogonal projector for $S^\perp$ is $I - P$.