Math 2270-3
Spring 2005

Explanation of the Pseudo-inverse

A lot of you had trouble with the problems from Section 5.4. This should help clarify the *why* for doing such problems.

One of the goals of this class is to understand how to solve a linear system $Ax = b$ for $x \in \mathbb{R}^m$, $b \in \mathbb{R}^n$. Two fundamental spaces that we have encountered are the image and the kernel of the n by m matrix A. Recall that the kernel is the set of all vectors $x \in \mathbb{R}^m$ where $Ax = 0$, and the image is the set of all vectors $y \in \mathbb{R}^n$ where $Ax = y$. What if we don't have a square matrix? Is it still possible to define a solution even when we don't have a consistent set of equations? These questions are at the heart of the matter.

Previous work had stated that if we can uniquely solve $Ax = b$ only if A is a square matrix (m=n) and rank of the augmented matrix is equal to n. However Chapter 5 expanded our vocabulary so we can talk about orthogonal complements, and in fact we have two key relations:

$$(image(A))^{\perp} = ker(A^T)$$
$$(ker(A))^{\perp} = image(A^T)$$

(The second one you showed in Problem 5.4.4) What a least squares solution does (for consistent systems) is project b onto the image of A and allows you solve the system $Ax = proj_{im(A)}(b)$. By forming the normal equations, the *least squares solution*, x^*, can be found by a multiplying b by $(A^TA)^{-1}A^T$. The matrix $(A^TA)^{-1}A^T$ is called the *pseudo-inverse*. The main point to take away from this is that a basis for $\mathbb{R}^n$ is found by the image of A and its orthogonal complement, the kernel of A^T.

The next question to ask is what about our solution space $x \in \mathbb{R}^m$? In Problem 5.4.4 you showed that the orthogonal complement to the kernel of A is the image of A^T. Problem 5.4.10 asked you to verify that a solution x to a consistent system $Ax = b$ has a part in the kernel of A and the image of A^T, and in fact the part of the solution in the image of A^T is unique (there is only one). So a *minimal solution* to the system $Ax = b$ is the one that lies entirely in the image of A^T. You can add anything to that minimal solution in the kernel and it will still solve your system. So now we have a relationship between our solution space (x) and our model space b):

$$\mathbb{R}^m : im(A^T) \bigoplus ker(A)$$
$$\mathbb{R}^n : im(A) \bigoplus ker(A^T)$$

Note that above, $Z = X \bigoplus Y$ just mean that the space Z is spanned by X and Y.
So the *minimal least squares solution* is the one that projects b onto the image of A and finds the unique vector x that lives the image of A^T. Exercise 11 and 13 built upon this, and you can show that this is a linear transformation. In other words, for any matrix A and b, you can find a vector x that is the minimal least squares solution to $Ax = b$, where $L^+(b) = x$. The matrix of this linear transformation is the *pseudo-inverse*

It might be helpful to ask if this result is consistent on what we already know. If we have an invertible matrix A, then we know that the kernel of A is zero, and the image of A will be all of $\mathbb{R}^n$.

Amazingly, this works for inconsistent linear systems as well. Look at part e of Problem 5.4.13. The matrix given has a row of zeros, so any vector with a non-zero second component will be an inconsistent system, previously unsolvable. By the concept of a pseudo-inverse allows us to discuss the "best" solution to a system, even when we can't solve it uniquely. Here "best" is a subjective term, but usually is synonymous with least squares.

All the above is a statement of a famous theorem:

Fredholm Alternative Theorem: Provided all the entries of A are real numbers, Tte equation $Ax = b$ has a solution if and only if $v^T b = 0$ for every vector v in $\ker(A^T)$.

Corollary: This solution is unique if and only if $\ker(A) = 0$

Notice that we have exactly what we are looking for: (a) when solutions to $Ax = b$ exist, and (b) when are they unique. The Fredholm Alternative theorem just projects b onto the image of A. If the solutions are unique, A is square and we can find the inverse. If they are not, then we can use the pseudo-inverse.

Neat, huh?