Sketch of Lecture 25

Example 119. (warmup) What is the pseudoinverse A^+ of $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$? Compute A^+A and AA^+.

Solution. The pseudoinverse is $A^+ = \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$.

The products are $A^+A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $AA^+ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Comment. Note how the pseudoinverse tries to behave like the regular inverse. But since A has only 2 columns, AA^+ can have rank at most 2, so it cannot be the full 3×3 identity.

Review.

- If the $m \times n$ matrix A has SVD $A = U\Sigma V^T$, then its pseudoinverse is $A^+ = V\Sigma^+ U^T$.

  Here, Σ^+, the pseudoinverse of Σ, is the $n \times m$ diagonal matrix, whose nonzero entries are the inverses of the entries of Σ.

- The system $A\boldsymbol{x} = \boldsymbol{b}$ has "optimal" solution $\boldsymbol{x} = A^+\boldsymbol{b}$.

  Here, "optimal" means that $\boldsymbol{x}$ is the smallest least squares solution. In particular:

  - If $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution, then $\boldsymbol{x} = A^+\boldsymbol{b}$ is that solution.
  - If $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least squares solution, then $\boldsymbol{x} = A^+\boldsymbol{b}$ is that least squares solution.
  - If $A\boldsymbol{x} = \boldsymbol{b}$ has many (possibly least squares) solutions, then $\boldsymbol{x} = A^+\boldsymbol{b}$ is one of these, namely the one with smallest norm.

    We haven't yet seen this case in action. The next examples illustrate the simplest case.

Example 120. Find the smallest solution to $A\boldsymbol{x} = [\,6\,]$ with $A = [\,2 \;\; 0 \;\; 0\,]$ in two ways: directly (because the equation is so simple) and using the pseudoinverse of $A = [\,2 \;\; 0 \;\; 0\,]$.

Solution. (direct) The general solution is $\boldsymbol{x} = \begin{bmatrix} 3 \\ s_1 \\ s_2 \end{bmatrix}$. Obviously, the smallest among these is $\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$.

Solution. $A^TA = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}[\,2 \;\; 0 \;\; 0\,] = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ has 4-eigenvector $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and 0-eigenvectors $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

$\boldsymbol{u}_1 = \frac{1}{\sigma_1}A\boldsymbol{v}_1 = \frac{1}{2}[\,2 \;\; 0 \;\; 0\,]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1$

Hence, $A = U\Sigma V^T$ with $U = [\,1\,]$, $\Sigma = [\,2 \;\; 0 \;\; 0\,]$, $V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

[Dang! We could (should?!) have realized from the beginning that A is already diagonal!]

The pseudoinverse is $A^+ = V\Sigma^+U^T = \begin{bmatrix} 1/2 \\ 0 \\ 0 \end{bmatrix}$ (which is immediate if we realize that A is diagonal).

Hence, the smallest solution is $\boldsymbol{x} = A^+[\,6\,] = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

Example 121.

(a) Find the pseudoinverse of $A = [\ 1\ \ 2\ \ 3\ ]$.

(b) Find the smallest solution to $x_1 + 2x_2 + 3x_3 = 6$.

As before, smallest solutions means the solution $\boldsymbol{x}$ such that $\|\boldsymbol{x}\|$ is as small as possible. One obvious solution is $[\ 1, 1, 1\ ]^T$, but is it the smallest?

Solution.

(a) $A^T A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [\ 1\ \ 2\ \ 3\ ] = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$ has 14-eigenvector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and 0-eigenvectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$.

$\boldsymbol{u}_1 = \frac{1}{\sigma_1} A\boldsymbol{v}_1 = \frac{1}{\sqrt{14}} [\ 1\ \ 2\ \ 3\ ] \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1$

Hence, $A = U\Sigma V^T$ with $U = [\ 1\ ]$, $\Sigma = [\ \sqrt{14}\ \ 0\ \ 0\ ]$, $V = \begin{bmatrix} 1/\sqrt{14} & * & * \\ 2/\sqrt{14} & * & * \\ 3/\sqrt{14} & * & * \end{bmatrix}$.

$A^+ = V\Sigma^+ U^T = \begin{bmatrix} 1/\sqrt{14} & * & * \\ 2/\sqrt{14} & * & * \\ 3/\sqrt{14} & * & * \end{bmatrix} [\ 1/\sqrt{14}\ \ 0\ \ 0\ ][\ 1\ ] = \frac{1}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Comment. No surprise on U. The only options for U are $U = [\ 1\ ]$ and $U = [\ -1\ ]$.

Comment. Realizing what we did here allows us to write down A^+ immediately for all $1 \times n$ matrices A. See Example 122.

Homework. Complete the SVD of A. That is, find an option for the two missing columns of V, so that V is an orthogonal matrix. In other words, find an orthonormal basis for the orthogonal complement of $\boldsymbol{v}_1$.

(b) We are solving $A\boldsymbol{x} = [\ 6\ ]$ with $A = [\ 1\ \ 2\ \ 3\ ]$ as in the previous example.

We conclude that the smallest solution is $\boldsymbol{x} = A^+ [\ 6\ ] = \frac{3}{7} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

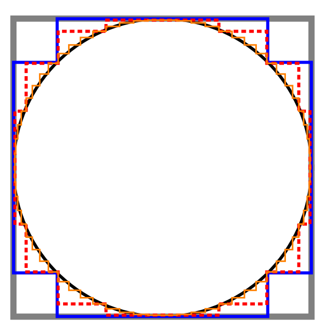
Compare. $\left\| \frac{3}{7} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\| = \frac{3}{7}\sqrt{14} \approx 1.604$ is indeed smaller than, say, $\left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\| = \sqrt{3} \approx 1.732$.

Geometric picture. The equation $x_1 + 2x_2 + 3x_3 = 6$ describes a plane (not through the origin), and we are asking for the point on that plane which is closest to the origin. That's a typical question in Calculus III. Try and use this geometric picture to solve the problem. Then compare with our earlier answer.

Example 122. More generally, find the pseudoinverse of $A = [\ a_1\ \ a_2\ \ a_3\ ]$.

Solution. Going through the previous example, we see that the answer will be $A^+ = \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|^2}$ with $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$.

Remark 123. (Pi Day!) π is the perimeter of a circle enclosed in a square with edge length 1. The perimeter of the square is 4, which approximates π. To get a better approximation, we "fold" the vertices of the square towards the circle (and get the blue polygon). This construction can be repeated for even better approximations and, in the limit, our shape will converge to the true circle. At each step, the perimeter is 4, so we conclude that $\pi = 4$, contrary to popular belief.



Can you pin-point the fallacy in this argument?

Comment. We'll actually come back to this. It's related to linear algebra in infinite dimensions.

Armin Straub
straub@southalabama.edu