Solutions for Homework Assignment 6  
Math/Comp 128 - Math 250NLA

1. Let $A = U\Sigma V^*$ be the SVD of $A$. We want to find the solution to the least squares problem

$$\min_x \|Ax - b\|_2,$$

with the smallest 2-norm. Using the SVD we can rewrite the least squares problem as follows

$$\|Ax - b\|_2 = \|U\Sigma V^* x - b\|_2 = \|\Sigma V^* x - U^* b\|_2 = \|\Sigma y - \tilde{b}\|_2.$$

We see that this system has an infinite number of solutions of the type

$$(y_1, y_2, \ldots, y_r, y_{r+1}, \ldots, y_n) = \left(\frac{\tilde{b}_1}{\sigma_1}, \frac{\tilde{b}_2}{\sigma_2}, \ldots, \frac{\tilde{b}_r}{\sigma_r}, y_{r+1}, \ldots, y_n\right),$$

where $(y_{r+1}, \ldots, y_n)$ can take any values. We want $x$ such that it has the smallest 2-norm. Since $V$ is a unitary matrix, we have that

$$\|x\|_2 = \|V^* x\|_2 = \|y\|_2.$$ 

So we give the values zero to $y_i$ for $i = r+1, \ldots, n$ which gives us the smallest 2-norm. To obtain $x$, we compute $x = V y = \sum_{i=1}^r y_i v_i = \sum_{i=1}^r \frac{u_i^T b}{\sigma_i} v_i = V(:, 1 : r) \Sigma^{-1} U(:, 1 : r^*)$.

2. This algorithm computes an approximate pseudoinverse of the matrix $A$.

3. We compute both roots using the formula

$$-b + \sqrt{b^2 - 4ac}$$

and

$$-b - \sqrt{b^2 - 4ac}.$$ 

When we use the latter formula we subtract two very close numbers, $-b$ and $\sqrt{b^2 - 4ac}$, and we lose important information because of rounding errors. This lost of information is reflected in the relative error.

4. Assume that the base $\beta$ is 2. Then, the machine epsilon is given by $\epsilon_{\text{machine}} = \frac{1}{2} 2^{1-t} = 2^t$ (formula 13.3 in the text). Recall that the machine epsilon is the smallest number which when added to 1 gives something different than 1. The algorithm starts with $E=1$. After $t$ steps, $E=\epsilon_{\text{machine}}$. Then $\text{EP1} = 1 + \epsilon_{\text{machine}} > 1$, so the algorithm enters the loop and computes $E=\epsilon_{\text{machine}}/2$. Now we have $\text{EP1} = 1 + \epsilon_{\text{machine}}/2 = 1$, then the algorithm stops and the last return is $E=\epsilon_{\text{machine}}/2$.

5.a. Assume that $A$ is a square matrix. If we have $x_i = x_j$ for some data points, then we will have two equal rows in the matrix $A$. Therefore, the matrix $A$ will be rank deficient.

5.b. By theorem 16.3. in the text, we know that

$$\frac{\|\hat{x} - x\|}{\|x\|} = O(k(A) \epsilon_{\text{machine}}).$$

If the size of the matrix $A$ is large, the condition number of $A$ is large. Then, by the theorem above, we could obtain a bad approximation.