Using SVD for least squares

Recall that a singular value decomposition is given by:

$$A = \begin{bmatrix} \vdots & & \vdots \\ u_1 & \cdots & u_m \\ \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & & 0 \\ 0 & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ 0 & & & & \cdots & 0 \end{bmatrix} \begin{bmatrix} \cdots & v_1^T & \cdots \\ & \vdots & \\ \cdots & v_n^T & \cdots \end{bmatrix}$$

where σ_i are the singular values.

Now back to our system of equations:

$$\underset{m\times n}{A} x = b$$

Assume that A has rank k (and hence k nonzero singular values σ_i) and recall that we want to minimize

$$\|r\|_2^2 = \|b - Ax\|_2^2 .$$

Substituting the SVD of A we find that

$$\|r\|_2^2 = \|b - Ax\|_2^2 = \|b - USV^T x\|_2^2$$

Where U and V are orthogonal and S is diagonal with k nonzero singular values

$$\|b - USV^T x\|_2^2 = \|U^T b - U^T U S V^T x\|_2^2 = \|U^T b - SV^T x\|_2^2$$

let $c = U^T b$ and $y = V^T x$ (and hence $x = Vy$) in $\|U^T b - SV^T x\|_2^2$. We now have

$$\|r\|_2^2 = \|c - Sy\|_2^2 .$$

Since S has only k nonzero diagonal elements, we have

$$\|r\|_2^2 = \sum_{i=1}^{k} (c_i - \sigma_i y_i)^2 + \sum_{i=k+1}^{n} c_i^2$$

which is minimized when $y_i = \frac{c_i}{\sigma_i}$ for $1 \le i \le k$.

<u>Thm</u>: let $A_{m\times n}$ matrix of rank k and let $A = USV^T$. The least squares solution of the system $Ax=b$ is

$$x = \sum_{i=1}^{k} (\sigma_i^{-1} c_i)\, v_i$$

where $c_i = u_i^T b$.

<u>Note</u> (1) if $m=n \Rightarrow A_{n\times n}$ and A is full column rank then $A^{-1}Ax = A^{-1}b \Rightarrow x = A^{-1}b$

however there are many situations where the inverse of A doesn't exist. In these cases we will approximate the inverse via the SVD which can turn a singular problem into a nonsingular one. We can also use the normal equations; i.e.

$$A^TAx = A^Tb \Rightarrow x = (A^TA)^{-1}A^Tb.$$

Recall that the Moore-Penrose pseudoinverse equals

$A^+ = (A^TA)^{-1}A^T$ if A is full rank. Furthermore we can use the SVD to compute $A^+ = VS^+U^T$.

where $S^+ = \begin{cases} 1/\sigma_i & i=1,\dots,k \\ 0 & \text{otherwise.} \end{cases}$

Note that even if A is ill-conditioned or singular SVD can give us a workable solution.

Example (1) [Equal number of equations and unknowns)

let $A=\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$ $\quad b=\begin{bmatrix}10\\ 5\end{bmatrix}$

A is square but not symmetric and singular ($|A|=0$)

$\Rightarrow A^{-1}$ doesn't exist, however Using the SVD we can approximate the inverse. The SVD approach tells us to compute eigenvalues and eigenvectors from

$$A^TA=\begin{bmatrix}5 & 5\\ 5 & 5\end{bmatrix} \quad \text{and} \quad AA^T=\begin{bmatrix}2 & 4\\ 4 & 8\end{bmatrix}$$

The eigenvalues from these matrices are $\lambda_1=0$ and $\lambda_2=10$.
Consequently the singular values of A are $\sigma_1=0$ and $\sigma_2=\sqrt{10}$
Therefore the rank of A is 1. The decomposition is then expressed as

$$A=USV^T=\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & \sqrt{10}\end{bmatrix}\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}^T$$

$$\Rightarrow A^{+}=VS^{+}U^T=\begin{bmatrix}1/\sqrt{2} & +1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & 1/\sqrt{10}\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$$

$$\Rightarrow X=\begin{bmatrix}1/10 & 1/5\\ 1/10 & 1/5\end{bmatrix}\begin{bmatrix}10\\ 5\end{bmatrix}=\begin{bmatrix}2\\ 2\end{bmatrix}$$

Example 2: [Overdetermined – more equations than unknowns].

let $C=\begin{bmatrix}1 & 1\\ 1 & 1\\ 0 & 0\end{bmatrix}$ $\quad b=\begin{bmatrix}b_1\\ b_2\\ b_3\end{bmatrix}$

$$\Rightarrow C^TC=\begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix} \quad \text{and} \quad CC^T=\begin{bmatrix}2 & 2 & 0\\ 2 & 2 & 0\\ 0 & 0 & 0\end{bmatrix}$$

The eigenvalues from $C'C$ are $\lambda_1 = 4$ and $\lambda_2 = 0$.

The eigenvalues from CC^T are $\lambda_1 = 4$, $\lambda_2 = 0$ and $\lambda_3 = 0$.

Consequently the ~~nonzero~~ singular values are $\sigma_1 = \sqrt{4}$ and $\sigma_2 = 0 \Rightarrow$ therefore the rank of C is 1. The SVD is

$$C = USV^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{4} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}^T$$

$$C^+ = VS^+U^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{4} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow x = C^+ b = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 + b_2 \\ b_1 + b_2 \end{bmatrix}.$$