The singular value decomposition for a real $m \times d$ matrix A

1. Let

$$\underbrace{U^T}_{d\times d}\underbrace{A^TA}_{d\times d}\underbrace{U}_{d\times d} = \Sigma^2 = \begin{bmatrix} \overset{r\times r}{\Sigma_1^2} & 0 \\ 0 & 0 \end{bmatrix} \quad (1)$$

where $\Sigma_1^2 = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_r^2 \end{bmatrix}$

i.e. $\sigma_1^2 \geq \sigma_2^2 \quad \geq \sigma_r^2 > 0$ are the positive eigenvalues of A^TA (the remaining $d-r$ eigenvalues are 0).

U is an orthogonal matrix

$$U = [\underline{u}_1 \underline{u}_2 \cdots \underline{u}_d]$$

$\underline{u}_i$ is an eigen~~value~~ vector for the i'th eigenvalue of A^TA. These eigenvectors are orthonormal.

2' Set $U_1 = [\underline{u}_1, \underline{u}_2 \; \underline{u}_r]$ a $d \times r$ matrix with orthonormal columns, and $U_2 = [\underline{u}_{r+1} \cdot \underline{u}_d]$ a $d \times (d-r)$ matrix with orthonormal columns.

$$U = [U_1 \; U_2]$$

3 It is easy to check from (1) that

$$U_1^T A^T A U_1 = \Sigma_1^2$$

and since Σ_1 is invertible it follows that

$$\underbrace{\boxed{\Sigma_1^{-1} U_1^T A^T}}_{V_1^T} \underbrace{\boxed{A U_1 \Sigma_1^{-1}}}_{\overset{\text{define}}{=} V_1} = I_{r \times r} \qquad (2)$$

V_1 is an $n \times r$ matrix, and (2) asserts that it has orthonormal columns

$$V_1 = [\underline{v}_1 \, \underline{v}_2 \;\; \underline{v}_r]$$

(therefore it also follows that $r \leq m$)

4 Using Gram-Schmidt find vectors $\underline{v}_{r+1}\ \underline{v}_{r+2} \cdots \underline{v}_m$ so that

$$\{\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_m\}$$

forms an orthonormal basis for $\mathbb{R}^m$

and define

$$V = [V_1\ V_2]$$

with $V_2 = [\underline{v}_{r+1} \cdots \underline{v}_m]$ (3)

5

We now calculate

$$V^T A U = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} A\, [U_1, U_2]$$

$$= \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} [AU_1, AU_2]$$

(cont) =
$$\begin{bmatrix} V_1^T A U_1 & V_1^T A U_2 \\ V_2^T A U_1 & V_2^T A U_2 \end{bmatrix}$$

Since $\underline{u}_{r+1} \cdots \underline{u}_d$ are eigenvectors of A^TA corresponding to eigenvalue 0, it follows that $A^TA\,\underline{u}_j = 0$ $j = r+1, \ldots, d$. Therefore $A\underline{u}_j = 0$ for $j = r+1, \ldots, d$. (remember $N(A^TA) = N(A)$). In other words $AU_2 = 0$ (matrix).

From the definition of V_1 (2) we get that the columns of AU_1 are multiples of the columns of V_1. From the definition of V_2 we have that the columns of V_2 are orthogonal to those of V_1

In other words the columns of AU_1 are orthogonal to the columns of V_2, or

$$V_2^T A U_1 = 0 \quad (\text{matrix})$$

6 We conclude that

$$V^T A U = \begin{bmatrix} \overbrace{V_1^T A U_1}^{r \times r} & 0 \\ 0 & 0 \end{bmatrix}$$

From (2) we see that

$$V_1^T A U_1 \Sigma_1^{-1} = I_{r \times r}$$

so that

$$\boxed{V_1^T A U_1 = \Sigma_1}$$

7 In conclusion

$$V^T A U = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\Sigma_1 = \begin{pmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \sigma_r \end{pmatrix}, \quad \sigma_i > 0, \quad 1 \le i \le r$$

and V and U are orthogonal matrices (of dimension $n \times n$ and $d \times d$, respectively).