Slide 26-Singular Value Decomposition

MAT2040 Linear Algebra

Motivation

Recall: If A is a real symmetric matrix, $A \in \mathbb{R}^{n \times n}$, we know that there is an orthogonal matrix Q such that $Q^{-1}AQ = Q^TAQ = \Lambda$, where Λ is a diagonal matrix. Thus, $A = Q\Lambda Q^T$ is the eigen decomposition.

Question: If $A \in \mathbb{R}^{m \times n}$, do we still have a similar matrix decomposition?

Yes. The idea is to use the A^TA or AA^T to do the eigen decomposition.

Singular Value Decomposition

For any $A \in \mathbb{R}^{m \times n}$, it can be decomposed into

$$A = U \Sigma V^T$$

where U is a $m \times m$ orthogonal matrix, V is a $n \times n$ orthogonal matrix, Σ is a diagonal-like matrix. $\Sigma = (\tilde{\sigma}_{ij})_{m \times n}$ is defined as $\tilde{\sigma}_{ij} = 0, \text{if} \, i \neq j$, $\tilde{\sigma}_{ii} = \sigma_i, i = 1, \cdots, \min(m, n)$.

Case 1: If $m \geq n$ (tall matrix), $\Sigma = (\tilde{\sigma}_{ij})_{m \times n}$ is defined as $\tilde{\sigma}_{ij} = 0, \text{if} \, i \neq j$, $\tilde{\sigma}_{ii} = \sigma_i, i = 1, \cdots, n$, where $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$. If rank of A is r, then $\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n = 0$.

Case 2: If $m < n$ (fat matrix), $\Sigma = (\tilde{\sigma}_{ij})_{m \times n}$ is defined as $\tilde{\sigma}_{ij} = 0, \text{if} \, i \neq j$, $\tilde{\sigma}_{ii} = \sigma_i, i = 1, \cdots, m$, where $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$. If rank of A is r, then $\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_m = 0$.

In fact, if $m \geq n$ and rank(A)=r, then

$$\Sigma = \left[\begin{array}{cccc|cccc} \sigma_1 & & & & & & & \\ & \sigma_2 & & & & & & \\ & & \ddots & & & & & \\ & & & \sigma_r & & & & \\ \hline & & & & \sigma_{r+1} & & & \\ & & & & & \ddots & & \\ & & & & & & \sigma_n & \\ & & & & & & & \end{array}\right]_{m\times n} = \begin{bmatrix} \Sigma_1 & O_1 \\ O_2 & O_3 \end{bmatrix}$$

where

$$\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n = 0$$

$\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, $O_1 = O_{r\times(n-r)}$, $O_2 = O_{(m-r)\times r}$, $O_3 = O_{(m-r)\times(n-r)}$

If $m < n$ and rank(A)=r, then

$$\Sigma = \left[\begin{array}{cccc|ccc} \sigma_1 & & & & & & \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r & & & \\ \hline & & & & \sigma_{r+1} & & \\ & & & & & \ddots & \\ & & & & & & \sigma_m \end{array}\right]_{m\times n} = \begin{bmatrix} \Sigma_1 & O_1 \\ O_2 & O_3 \end{bmatrix}$$

where

$$\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_m = 0$$

$\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, $O_1 = O_{r\times(n-r)}$, $O_2 = O_{(m-r)\times r}$, $O_3 = O_{(m-r)\times(n-r)}$

If $rank(A) = r$, then for both cases ($m \geq n$ and $m \leq n$), Σ is defined as

$$\Sigma_{m\times n} = \left[\begin{array}{cc} \Sigma_1 & O_1 \\ O_2 & O_3 \end{array} \right]$$

$\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r), O_1 = O_{r\times(n-r)}, O_2 = O_{(m-r)\times r}, O_3 = O_{(m-r)\times(n-r)}$, where $\sigma_1, \cdots, \sigma_r$ are positive numbers.

Theorem 25.1 Let A be a $m \times n$ real matrix, then A has the singular value decomposition $A = U\Sigma V^T$.

Analysis:

$A = U\Sigma V^T \Rightarrow A^T = V\Sigma^T U^T \Rightarrow AA^T = U\Sigma\Sigma^T U^T$ and $A^T A = V\Sigma^T\Sigma V^T \Rightarrow U^{-1}AA^T U = \Sigma\Sigma^T$ and $V^{-1}A^T AV = \Sigma^T\Sigma$.

Suppose that $r(A) = r$, then

$$\Sigma = \begin{bmatrix} \Sigma_1 & O_1 \\ O_2 & O_3 \end{bmatrix}$$

where $\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r), \sigma_1 \geq \cdots \geq \sigma_r > 0$, $O_1 = O_{r\times(n-r)}$, $O_2 = O_{(m-r)\times r}$, $O_3 = O_{(m-r)\times(n-r)}$.

$(\Sigma^T\Sigma)_{n\times n} = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots, 0)$ (with $n-r$ zeros elements on the diagonal). Thus, $\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots, 0$ (with $n-r$ zeros) are eigenvalues of $A^T A$.

$(\Sigma\Sigma^T)_{m\times m} = \mathrm{diag}(\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots, 0)$ (with $m-r$ zeros elements on the diagonal). Thus, $\sigma_1^2, \cdots, \sigma_r^2, 0, \cdots, 0$ (with $m-r$ zeros) are eigenvalues of AA^T.

Proof. Without loss of generality, we first consider $m \geq n$. The case for $m < n$ can be proved in a similar way.

Since $A^T A$ is a $n \times n$ real symmetric matrix, which is diagonalizable by spectral theorem. All eigenvalues of $A^T A$ are nonnegative. (Suppose that $A^T A\mathbf{x} = \lambda\mathbf{x}(\mathbf{x} \neq \mathbf{0})$, then $\mathbf{x}^T A^T A\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x}$, thus $\lambda = \frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2} \geq 0$.)

The construction for $A = U\Sigma V^T$ is as follows:

Step 1 (construction of V). Suppose $rank(A) = r$, then $rank(A^T A) = r$. Since $A^T A$ is symmetric, there is an orthogonal matrix V that diagonalizes matrix $A^T A$ ($V^T A^T AV = \Lambda$), and the rank of $A^T A$ also equals to the number of nonzero eigenvalues of $A^T A$ (rank($A^T A$)=rank(Λ)=the number of nonzero eigenvalues of $A^T A$).

Suppose that $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$ be the eigenvalues of $A^T A$. The singular values of A are defined as $\sigma_i = \sqrt{\lambda_i}, i = 1, \cdots, n$. Then $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0$.

Let $V \triangleq [\mathbf{v}_1, \cdots, \mathbf{v}_n]$, where $\mathbf{v}_1, \cdots, \mathbf{v}_n$ are the eigenvectors of $A^T A$ corresponds to eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$, respectively. Since V is an orthogonal matrix, $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is the orthonormal set.

Let $V = [V_1, V_2]$, where $V_1 = [\mathbf{v}_1, \cdots, \mathbf{v}_r]$, $V_2 = [\mathbf{v}_{r+1}, \cdots, \mathbf{v}_n]$

Since $A^T A\mathbf{v}_i = \mathbf{0}, \quad i = r+1, \cdots, n$ and Null($A^T A$)=Null(A). Thus, $A\mathbf{v}_i = \mathbf{0}, \quad i = r+1, \cdots, n$ and $AV_2 = O$.

Since V is an orthogonal matrix, one has

$$I = VV^T = [V_1, V_2][V_1, V_2]^T = V_1 V_1^T + V_2 V_2^T$$

$$A = AI = A(V_1 V_1^T + V_2 V_2^T) = AV_1 V_1^T + AV_2 V_2^T = AV_1 V_1^T$$

Step 2 (construction of Σ). As discussed in the previous slides, once the singular values are obtained, one can construct Σ as follows: let $\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r), O_1 = O_{r\times(n-r)}, O_2 = O_{(m-r)\times r}, O_3 = O_{(m-r)\times(n-r)}$, then define

$$\Sigma = \begin{bmatrix} \Sigma_1 & O_1 \\ O_2 & O_3 \end{bmatrix}.$$

Step 3 (construction of U). To complete the proof, we need to construct U, we need to find the $m \times m$ orthogonal matrix U such that $A = U\Sigma V^T$. This gives $AV = U\Sigma$. Comparing the first r columns of this identity, one has $A\mathbf{v}_i = \sigma_i \mathbf{u}_i = \sqrt{\lambda_i}\mathbf{u}_i, i = 1, \cdots, r$. Define $\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}} A\mathbf{v}_i, i = 1, \cdots, r$, it follows that $AV_1 = U_1\Sigma_1$, where $U_1 = [\mathbf{u}_1, \cdots, \mathbf{u}_r]$.

Now we have $A^T\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}} A^T A\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}}\lambda_i \mathbf{v}_i$ $(i = 1, \cdots, r)$, thus, $\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}} A^T\mathbf{u}_i$ $(i = 1, \cdots, r)$.

In addition, $AA^T\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}}A(A^TA\mathbf{v}_i) = \frac{1}{\sqrt{\lambda_i}}\lambda_i A\mathbf{v}_i = \lambda_i\frac{1}{\sqrt{\lambda_i}}A\mathbf{v}_i = \lambda_i\mathbf{u}_i$ since $A^TA\mathbf{v}_i = \lambda_i\mathbf{v}_i$. Thus, λ_i $(i = 1, \cdots, r)$ are the nonzero eigenvalues of AA^T and $\mathbf{u}_i$ $(i = 1, \cdots, r)$ are the corresponding eigenvectors.

Moreover, $\mathbf{u}_i^T\mathbf{u}_j = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}}\mathbf{v}_i^TA^TA\mathbf{v}_j = \frac{\sqrt{\lambda_j}}{\sqrt{\lambda_i}}\mathbf{v}_i^T\mathbf{v}_j = \delta_{ij}$ $(i, j = 1, \cdots, r)$, where $A^TA\mathbf{v}_j = \lambda_j\mathbf{v}_j$ is used. Here $\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$

Thus, $\{\mathbf{u_1}, \cdots, \mathbf{u_r}\}$ is the orthonormal set. In addition, since $\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}}A\mathbf{v}_i, i = 1, \cdots, r$, $\{\mathbf{u_1}, \cdots, \mathbf{u_r}\} \in \mathrm{Col}(A)$. In addition, $\dim(\mathrm{Col}(A)) = rank(A) = r$, thus, $\{\mathbf{u_1}, \cdots, \mathbf{u_r}\}$ is an orthonormal basis of $\mathrm{Col}(A)$.

Since $\mathrm{Col}(A)^\perp = \mathrm{Null}(A^T)$, thus $\dim(\mathrm{Null}(A^T)) = m - r$ since $\mathrm{Col}(A)$ is a subspace of $\mathbb{R}^m$, $rank(A) = r$ and $\dim(\mathrm{Col}(A)) + \dim(\mathrm{Col}(A)^\perp) = m$.

Let $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$ are orthonormal basis of Null(A^T), then $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$ satisfies $A^T \mathbf{u}_i = \mathbf{0}, i = r+1, \cdots, m$. Indeed, $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$ are also the orthonormal basis of Null(AA^T) since Null(AA^T)=Null(A^T).

Set $U_2 = [\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m]$ and $U = [U_1, U_2]$, then $\{\mathbf{u_1}, \cdots, \mathbf{u_r}, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_m\}$ is an orthonormal basis of $\mathbb{R}^m$. By using the theorem 19.19.

Step 4 (Verification of $A = U\Sigma V^T$). Compute

$$U\Sigma V^T = [U_1, U_2] \left[\begin{array}{cc} \Sigma_1 & O_1 \\ O_2 & O_3 \end{array} \right] [V_1, V_2]^T = U_1 \Sigma_1 V_1^T = AV_1 V_1^T = A$$

Thus A can be written as:

$$A = U\Sigma V^T$$

This is called **Singular Value Decomposition**.

Example 25.2 Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$$

then

$$A^T A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

whose eigenvalues are $\lambda_1 = 4, \lambda_2 = 0$. The unit eigevector w.r.t. $\lambda_1 = 4$ is

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \left[\begin{array}{c} 1 \\ 1 \end{array} \right]$$

The unit eigevector w.r.t. $\lambda_2 = 0$ is

$$\mathbf{v}_2 = \frac{1}{\sqrt{2}} \left[\begin{array}{c} 1 \\ -1 \end{array} \right]$$

Now $\sigma_1 = \sqrt{\lambda_1} = 2$ so

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \left[\begin{array}{c} 1 \\ 1 \end{array} \right] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

is the eigenvector of AA^T corresponding to eigenvalue $\lambda_1 = 4$.
In addition,

$$A^T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

An orthonormal basis for Null(A^T)=Null(AA^T) is

$$\{\mathbf{u}_2, \mathbf{u}_3\} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} = A = U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Note: For real symmetric matrix S, the rank of matrix S equals to the number of nonzero eigenvalues. Since S is real symmetric, there is an orthogonal matrix Q, such that $Q^TSQ = \Lambda$ and rank(S)=rank(Q^TSQ)=rank(Λ)=the number of nonzero eigenvalues of S.

Remark 1

For $A \in \mathbb{R}^{m\times n}$, $A = U\Sigma V^T$ and rank(A)=r, one has
$\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$ are the eigenvalues of A^TA.
$\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_m = 0$ are the eigenvalues of AA^T.

(1) If $m \geq n$, as shown in above, let $\sigma_i = \sqrt{\lambda_i}(i = 1, \cdots, n)$, then $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0$ (with $n - r$ zeros) are the **singular values**.
If $m < n$, let $\sigma_i = \sqrt{\lambda_i}(i = 1, \cdots, m)$, then $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_m = 0$ (with $m - r$ zeros) are the **singular values**.

(2) The singular values of A are unique, but the orthogonal matrices U and V are not unique.

(3) $AV = U\Sigma$, Thus, $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$, $\sigma_i = \sqrt{\lambda_i} > 0$ $(i = 1, \cdots, r)$ and $A\mathbf{v}_i = \mathbf{0}$, $(i = r+1, \cdots, n)$.

(4) Take transpose for $A = U\Sigma V^T$, thus $A^T = V\Sigma^T U^T$. $A^T U = V\Sigma^T$, write in the vector form $A^T \mathbf{u}_j = \sigma_j \mathbf{v}_j$ $(j = 1, \cdots, r)$, $A^T \mathbf{u}_j = \mathbf{0}$ $(j = r+1, \cdots, m)$.

And the columns of U are called **left singular vector** of A; the columns of V are called **right singular vector** of A;

Remark 2 The rank of $m \times n$ matrix A is the number of nonzero singular values.

• The number of nonzero singular values (counting the multiplicity) equals to the rank of A.

• rank(A) $\neq$ number of nonzero eigenvalues. Example: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ then eigenvalues are $\lambda_1 = \lambda_2 = 0$, the number of nonzero eigenvalue is 0, but rank$(A) = 1$.

• rank(A) $=$ number of nonzero eigenvalues if A is real symmetric.

Remark 3 Four fundamental subspaces

$$
\begin{aligned}
A =& U\Sigma V^T \\
=& [\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_m] \begin{bmatrix} \Sigma_1 & O_1 \\ O_2 & O_3 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \\ \mathbf{v}_{r+1}^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}
\end{aligned}
$$

where $O_1 = O_{r\times(n-r)}, O_2 = O_{(m-r)\times r}, O_3 = O_{(m-r)\times(n-r)}$

1) First r columns of V is an orthonormal basis for Row(A)=Col(A^T)=$(Null(A))^{\perp}$, since $A^TA\mathbf{v}_i = \lambda_i \mathbf{v}_i (i = 1, \cdots, r)$, $\lambda_1 \geq \cdots \geq \lambda_r > 0$ ($\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}} A^T \mathbf{u_i} (i = 1, \cdots, r)$), $\mathbf{v}_i (i = 1, \cdots, r)$ are orthonormal vector set and $dim(Row(A)) = r$.

2) Last $n - r$ columns of V is an orthonormal basis for Null(A), since $A^TA\mathbf{v}_i = 0 (i = r+1, \cdots, n)$, Null($A^TA$)= Null($A$), $\mathbf{v}_i (i = r+1, \cdots, n)$ are orthonormal vector set and $A\mathbf{v}_i = 0 (i = r+1, \cdots, n)$.

3) First r columns of U is an orthonormal basis for Col(A)=$(Null(A^T))^{\perp}$ since $\mathbf{u}_i = \frac{1}{\sqrt{\lambda_i}} A\mathbf{v}_i, i = 1, \cdots, r$, $\mathbf{u}_i (i = 1, \cdots, r)$ are orthonormal vector set and $dim(Col(A)) = r$.

4) Last $m - r$ columns of U is an orthonormal basis for Null(A^T), since $AA^T\mathbf{u}_i = 0 (i = r+1, \cdots, m)$, Null($AA^T$)= Null($A^T$), $\mathbf{u}_i (i = r+1, \cdots, m)$ are orthonormal vector set and $A^T\mathbf{u}_i = 0 (i = r+1, \cdots, m)$.

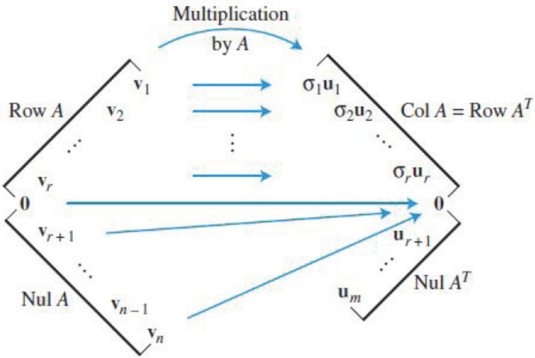

Figure: Here $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \cdots, r$

Remark 4 Compact SVD:

$$\begin{aligned} A =& U\Sigma V^T \\ =& [\mathbf{u}_1, \cdots, \mathbf{u}_r]\Sigma_1 \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix} \end{aligned}$$

where

$$\Sigma_1 = diag(\sigma_1, \cdots, \sigma_r)$$

This gives $A\mathbf{v}_i = \sigma_i \mathbf{u}_i, \sigma_i = \sqrt{\lambda_i}, (i = 1, \cdots, r)$ and

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

Remark: $\mathbf{uv}^T$ is the out product of two vectors $\mathbf{u}$ and $\mathbf{v}$.

The **outer product** $\mathbf{x}\mathbf{y}^T$ will result in a matrix.

$$\begin{aligned}
\mathbf{x}\mathbf{y}^T &= \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 \cdots & y_n \end{bmatrix} \\
&= \begin{bmatrix} x_1 y_1 & \cdots & x_1 y_n \\ x_2 y_1 & \cdots & x_2 y_n \\ \vdots & & \\ x_m y_1 & \cdots & x_m y_n \end{bmatrix} \\
&\triangleq [\mathbf{b}_1, \cdots, \mathbf{b}_n]
\end{aligned}$$

Suppose $\mathbf{x}, \mathbf{y}$ are both nonzero vectors. Let $\mathbf{y} = [y_1, \cdots, y_n]^T$, and assume $y_1 \neq 0$, then $\mathbf{b}_i = \frac{y_i}{y_1}\mathbf{b}_1, i = 2, \cdots, n$. Thus, Col($\mathbf{x}\mathbf{y}^T$)$=\mathrm{Span}(\mathbf{b}_1)$. Therefore, the rank of the outer product $\mathbf{x}\mathbf{y}^T$ is 1.

Proposition Every rank 1 matrix A has the form $A = \mathbf{x}\mathbf{y}^T =$ column vector$\times$ row vector.

Remark 5 Vector form

- Recall a real symmetric matrix A has the eigen decomposition as follows:

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

where λ_i $(i = 1, \cdots, n)$ are the eigenvalues of A, and $Q = [\mathbf{q}_1, \cdots, \mathbf{q}_n]$ is the orthogonal matrix which diagonalizes A.

- For $A \in \mathbb{R}^{m \times n}$, the SVD decomposition gives:

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

where $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \cdots, r$, $\lambda_i (i = 1, \cdots, r)$ are nonzero eigenvalues of $A^T A$ (or nonzero eigenvalues of AA^T) and $r = \text{rank}(A) =$ number of nonzero singular values of $A(\sigma_i(i = 1, \cdots, r)$ are nonzero singular values).