

# Linear Algebra

Rekha Santhanam

Johns Hopkins Univ.

May 3, 2009

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable,

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable, therefore there exists an orthonormal eigenbasis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ .

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable, therefore there exists an orthonormal eigenbasis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ .

Then  $\vec{x} = c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$  for some  $c_1, \dots, c_n$ .

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable, therefore there exists an orthonormal eigenbasis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ .

Then  $\vec{x} = c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$  for some  $c_1, \dots, c_n$ .

Therefore,

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable, therefore there exists an orthonormal eigenbasis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ .

Then  $\vec{x} = c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$  for some  $c_1, \dots, c_n$ .

Therefore,

$$\begin{aligned} q(\vec{x}) &= \vec{x}^T A \vec{x} \\ &= (c_1 \vec{u}_1^T + \dots + c_n \vec{u}_n^T) A (c_1 \vec{u}_1 + \dots + c_n \vec{u}_n) \end{aligned}$$

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable, therefore there exists an orthonormal eigenbasis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ .

Then  $\vec{x} = c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$  for some  $c_1, \dots, c_n$ .

Therefore,

$$\begin{aligned} q(\vec{x}) &= \vec{x}^T A \vec{x} \\ &= (c_1 \vec{u}_1^T + \dots + c_n \vec{u}_n^T) A (c_1 \vec{u}_1 + \dots + c_n \vec{u}_n) \\ &= (c_1 \vec{u}_1^T + \dots + c_n \vec{u}_n^T) (c_1 \lambda_1 \vec{u}_1 + \dots + c_n \lambda_n \vec{u}_n) \end{aligned}$$

A quadratic form  $q$  is a function on  $n$ -variables defined as

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

where  $A$  is a  $n \times n$  symmetric matrix.

Every symmetric matrix is orthogonally diagonalizable, therefore there exists an orthonormal eigenbasis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ .

Then  $\vec{x} = c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$  for some  $c_1, \dots, c_n$ .

Therefore,

$$\begin{aligned} q(\vec{x}) &= \vec{x}^T A \vec{x} \\ &= (c_1 \vec{u}_1^T + \dots + c_n \vec{u}_n^T) A (c_1 \vec{u}_1 + \dots + c_n \vec{u}_n) \\ &= (c_1 \vec{u}_1^T + \dots + c_n \vec{u}_n^T) (c_1 \lambda_1 \vec{u}_1 + \dots + c_n \lambda_n \vec{u}_n) \\ &= c_1^2 \lambda_1 + \dots + c_n^2 \lambda_n \end{aligned}$$

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite.

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ .

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ . Then the

$$\text{matrix } A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ . Then the

$$\text{matrix } A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

The matrix  $A$  has eigenvalues  $\frac{1}{2}$  and  $-\frac{1}{2}$ ,

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ . Then the

$$\text{matrix } A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

The matrix  $A$  has eigenvalues  $\frac{1}{2}$  and  $-\frac{1}{2}$ , with corresponding eigenvectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ . Then the

$$\text{matrix } A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

The matrix  $A$  has eigenvalues  $\frac{1}{2}$  and  $-\frac{1}{2}$ , with corresponding eigenvectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \text{ Then this gives a orthonormal eigenbasis for } \mathbb{R}^2 \text{ as } \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ . Then the

$$\text{matrix } A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

The matrix  $A$  has eigenvalues  $\frac{1}{2}$  and  $-\frac{1}{2}$ , with corresponding eigenvectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \text{ Then this gives an orthonormal eigenbasis for } \mathbb{R}^2 \text{ as } \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

Therefore, by previous computation we see that for  $c_1, c_2$  determined by  $x_1$  and  $x_2$  in terms of  $\vec{u}_1, \vec{u}_2$  we have,

If all the eigenvalues are positive (non-negative), the matrix  $A$  is said to be positive (semi)definite and if all the eigenvalues are negative (non-positive), the matrix is said to be negative (semi)definite. If some are positive, some are negative then the matrix is indefinite.

For example, if the quadratic form is defined as  $q(x_1, x_2) = x_1x_2$ . Then the

$$\text{matrix } A = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}.$$

The matrix  $A$  has eigenvalues  $\frac{1}{2}$  and  $-\frac{1}{2}$ , with corresponding eigenvectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \text{ Then this gives an orthonormal eigenbasis for } \mathbb{R}^2 \text{ as } \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

Therefore, by previous computation we see that for  $c_1, c_2$  determined by  $x_1$  and  $x_2$  in terms of  $\vec{u}_1, \vec{u}_2$  we have,

$$q(x_1, x_2) = \frac{1}{2}c_1^2 - \frac{1}{2}c_2^2.$$

In our example we see that level curves (curves of the form  $q(x_1, \dots, x_n) = k$ ) for the given quadratic forms will be hyperbolas with respect to the new coordinate basis.

Thus the knowledge of orthogonal diagonalization of  $A$  allows us to rewrite the quadratic form in a new set of coordinates with respect to which the quadratic form is easier to understand.

Now we will describe another application of symmetric matrices by using the idea that given any matrix  $A$  (does not even have to be a square matrix),  $A^T A$  is a symmetric matrix.

Remember the example that you solved in your homework which showed that an invertible transformation from  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  maps a unit circle onto an ellipse. This follows from the fact that an invertible transformation described by  $A$  maps some pair of orthogonal vectors on the unit circle onto orthogonal vectors.

In general, an orthogonal transformation takes orthogonal vectors to orthogonal vectors.

Remember the example that you solved in your homework which showed that an invertible transformation from  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  maps a unit circle onto an ellipse. This follows from the fact that an invertible transformation described by  $A$  maps some pair of orthogonal vectors on the unit circle onto orthogonal vectors.

In general, an orthogonal transformation takes orthogonal vectors to orthogonal vectors.

Given a transformation described by a symmetric matrix, the eigenvectors corresponding to distinct eigenvalues are orthogonal.

Remember the example that you solved in your homework which showed that an invertible transformation from  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  maps a unit circle onto an ellipse. This follows from the fact that an invertible transformation described by  $A$  maps some pair of orthogonal vectors on the unit circle onto orthogonal vectors.

In general, an orthogonal transformation takes orthogonal vectors to orthogonal vectors.

Given a transformation described by a symmetric matrix, the eigenvectors corresponding to distinct eigenvalues are orthogonal.

In fact, it turns out that given any linear transformation  $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$  we can find a pair of orthogonal vectors  $\vec{v}_1, \vec{v}_2$  such that  $L(\vec{v}_1)$  and  $L(\vec{v}_2)$  are orthogonal. Of course it is possible that  $L(\vec{v}_i)$  is actually the zero vector!

Remember the example that you solved in your homework which showed that an invertible transformation from  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  maps a unit circle onto an ellipse. This follows from the fact that an invertible transformation described by  $A$  maps some pair of orthogonal vectors on the unit circle onto orthogonal vectors.

In general, an orthogonal transformation takes orthogonal vectors to orthogonal vectors.

Given a transformation described by a symmetric matrix, the eigenvectors corresponding to distinct eigenvalues are orthogonal.

In fact, it turns out that given any linear transformation  $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$  we can find a pair of orthogonal vectors  $\vec{v}_1, \vec{v}_2$  such that  $L(\vec{v}_1)$  and  $L(\vec{v}_2)$  are orthogonal. Of course it is possible that  $L(\vec{v}_i)$  is actually the zero vector!

Given the matrix  $A$  representing this transformation,  $A^T A$  is symmetric.

Given the matrix  $A$  representing this transformation,  $A^T A$  is symmetric. Let  $\vec{v}_1$  and  $\vec{v}_2$  be orthogonal eigenvectors of  $A^T A$ .

Given the matrix  $A$  representing this transformation,  $A^T A$  is symmetric.

Let  $\vec{v}_1$  and  $\vec{v}_2$  be orthogonal eigenvectors of  $A^T A$ .

Then

$$(A\vec{v}_1)^T (A\vec{v}_2) = \vec{v}_1^T A^T A \vec{v}_2 = \lambda_2 \vec{v}_1^T \vec{v}_2 = 0.$$

Given the matrix  $A$  representing this transformation,  $A^T A$  is symmetric.

Let  $\vec{v}_1$  and  $\vec{v}_2$  be orthogonal eigenvectors of  $A^T A$ .

Then

$$(A\vec{v}_1)^T (A\vec{v}_2) = \vec{v}_1^T A^T A \vec{v}_2 = \lambda_2 \vec{v}_1^T \vec{v}_2 = 0.$$

This allows us to prove that any invertible linear transformation from  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  maps a circle into an ellipse.

Let  $A$  be a  $n \times m$  matrix.

Let  $A$  be a  $n \times m$  matrix.

Then  $A^T A$  is a symmetric  $m \times m$  matrix. Note that if  $\lambda$  is an eigenvalue of  $A^T A$  with a corresponding eigenvector  $\vec{v}$  then

$$\|A\vec{v}\|^2 = (A\vec{v})^T A\vec{v} = \vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v} = \lambda \|\vec{v}\|^2.$$

This implies that  $\lambda \geq 0$ . The eigenvalues of  $A^T A$  are all positive real. Define the singular values of  $A$  as the square roots of the eigenvalues of  $A^T A$ .

Let  $A$  be a  $n \times m$  matrix.

Then  $A^T A$  is a symmetric  $m \times m$  matrix. Note that if  $\lambda$  is an eigenvalue of  $A^T A$  with a corresponding eigenvector  $\vec{v}$  then

$$\|A\vec{v}\|^2 = (A\vec{v})^T A\vec{v} = \vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v} = \lambda \|\vec{v}\|^2.$$

This implies that  $\lambda \geq 0$ . The eigenvalues of  $A^T A$  are all positive real. Define the singular values of  $A$  as the square roots of the eigenvalues of  $A^T A$ .

Write these in a descending order of magnitudes, denoted by  $\sigma_1, \dots, \sigma_m$ .

Let  $A$  be a  $n \times m$  matrix.

Then  $A^T A$  is a symmetric  $m \times m$  matrix. Note that if  $\lambda$  is an eigenvalue of  $A^T A$  with a corresponding eigenvector  $\vec{v}$  then

$$\|A\vec{v}\|^2 = (A\vec{v})^T A\vec{v} = \vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v} = \lambda \|\vec{v}\|^2.$$

This implies that  $\lambda \geq 0$ . The eigenvalues of  $A^T A$  are all positive real. Define the singular values of  $A$  as the square roots of the eigenvalues of  $A^T A$ .

Write these in a descending order of magnitudes, denoted by  $\sigma_1, \dots, \sigma_m$ . Let  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^m$  be orthonormal eigenvectors of  $A^T A$  corresponding to  $\sigma_1^2, \dots, \sigma_m^2$ .

Let  $A$  be a  $n \times m$  matrix.

Then  $A^T A$  is a symmetric  $m \times m$  matrix. Note that if  $\lambda$  is an eigenvalue of  $A^T A$  with a corresponding eigenvector  $\vec{v}$  then

$$\|A\vec{v}\|^2 = (A\vec{v})^T A\vec{v} = \vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v} = \lambda \|\vec{v}\|^2.$$

This implies that  $\lambda \geq 0$ . The eigenvalues of  $A^T A$  are all positive real. Define the singular values of  $A$  as the square roots of the eigenvalues of  $A^T A$ .

Write these in a descending order of magnitudes, denoted by  $\sigma_1, \dots, \sigma_m$ . Let  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^m$  be orthonormal eigenvectors of  $A^T A$  corresponding to  $\sigma_1^2, \dots, \sigma_m^2$ .

Then by previous argument  $A\vec{v}_1, \dots, A\vec{v}_m$  are orthogonal to each other,

Let  $A$  be a  $n \times m$  matrix.

Then  $A^T A$  is a symmetric  $m \times m$  matrix. Note that if  $\lambda$  is an eigenvalue of  $A^T A$  with a corresponding eigenvector  $\vec{v}$  then

$$\|A\vec{v}\|^2 = (A\vec{v})^T A\vec{v} = \vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v} = \lambda \|\vec{v}\|^2.$$

This implies that  $\lambda \geq 0$ . The eigenvalues of  $A^T A$  are all positive real. Define the singular values of  $A$  as the square roots of the eigenvalues of  $A^T A$ .

Write these in a descending order of magnitudes, denoted by  $\sigma_1, \dots, \sigma_m$ . Let  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^m$  be orthonormal eigenvectors of  $A^T A$  corresponding to  $\sigma_1^2, \dots, \sigma_m^2$ .

Then by previous argument  $A\vec{v}_1, \dots, A\vec{v}_m$  are orthogonal to each other, (Note some of them may be zero since the transformation is not invertible)

Let  $A$  be a  $n \times m$  matrix.

Then  $A^T A$  is a symmetric  $m \times m$  matrix. Note that if  $\lambda$  is an eigenvalue of  $A^T A$  with a corresponding eigenvector  $\vec{v}$  then

$$\|A\vec{v}\|^2 = (A\vec{v})^T A\vec{v} = \vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v} = \lambda \|\vec{v}\|^2.$$

This implies that  $\lambda \geq 0$ . The eigenvalues of  $A^T A$  are all positive real. Define the singular values of  $A$  as the square roots of the eigenvalues of  $A^T A$ .

Write these in a descending order of magnitudes, denoted by  $\sigma_1, \dots, \sigma_m$ . Let  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^m$  be orthonormal eigenvectors of  $A^T A$  corresponding to  $\sigma_1^2, \dots, \sigma_m^2$ .

Then by previous argument  $A\vec{v}_1, \dots, A\vec{v}_m$  are orthogonal to each other, (Note some of them may be zero since the transformation is not invertible) such that  $\|A\vec{v}_i\| = \sigma_i$  for all  $i = 1, \dots, m$ .

This is because

$$(A\vec{v}_i)^T (A\vec{v}_i) = \vec{v}_i^T A^T A\vec{v}_i = \lambda_i \vec{v}_i^T \vec{v}_i = \lambda_i.$$

These vectors describe the image and kernel of the transformation  $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$  defined by  $L(\vec{v}) = A\vec{v}$ .

These vectors describe the image and kernel of the transformation  $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$  defined by  $L(\vec{v}) = A\vec{v}$ .

Let  $A$  have rank  $r$ . Then the set  $\{A\vec{v}_1, A\vec{v}_2, \dots, A\vec{v}_m\}$  can have at most  $r$  non-zero vectors, since being orthogonal implies the set is linearly independent.

Then  $A\vec{v}_i = \sigma_i \vec{v}_i$  and  $\sigma_i$ 's are in descending order implies that  $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_m = 0$ .

Therefore,  $A\vec{v}_{r+1} = \dots = A\vec{v}_m = \vec{0}$  implies that  $\{\vec{v}_{r+1}, \dots, \vec{v}_m\} \subset \text{Ker } A$ .

By rank-nullity theorem we know that  $\text{Dim Ker } A = m - r$ . Therefore the orthonormal set  $\{\vec{v}_{r+1}, \dots, \vec{v}_m\}$  is a basis of  $\text{Ker } A$ .

This implies that  $\{A\vec{v}_1, \dots, A\vec{v}_r\}$  is a orthogonal set of non-zero vectors (otherwise we would have a bigger linearly independent subset in  $\text{Ker } A$ ) and hence linearly independent.

Since  $\text{Rank of } A = r$  then  $\{A\vec{v}_1, \dots, A\vec{v}_r\}$  is a basis of Image of  $A$ .

Define

$$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i \in \mathbb{R}^n \implies \vec{u}_i \sigma_i = A \vec{v}_i$$

for all  $i = 1, \dots, r$ .

Define

$$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i \in \mathbb{R}^n \implies \vec{u}_i \sigma_i = A \vec{v}_i$$

for all  $i = 1, \dots, r$ .

This gives us a orthonormal subset of  $\mathbb{R}^n$ . We can extend  $\{\vec{u}_1, \dots, \vec{u}_r\}$  to an orthonormal basis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  of  $\mathbb{R}^n$ .

Define

$$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i \in \mathbb{R}^n \implies \vec{u}_i \sigma_i = A \vec{v}_i$$

for all  $i = 1, \dots, r$ .

This gives us a orthonormal subset of  $\mathbb{R}^n$ . We can extend  $\{\vec{u}_1, \dots, \vec{u}_r\}$  to an orthonormal basis  $\{\vec{u}_1, \dots, \vec{u}_n\}$  of  $\mathbb{R}^n$ .

Then we can write

$$A[\vec{v}_1 \cdots \vec{v}_m] = [\vec{u}_1 \cdots \vec{u}_n] \begin{bmatrix} \sigma_1 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \sigma_r & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}_{n \times m} .$$

This is called the singular value decomposition of  $A = U \Sigma V^T$ .

$$\text{Let } A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

$$\text{Then } A^T A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

The eigenvalues of  $A^T A$  are 0, 1, 3 and,  $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ ,  $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$ .

Then  $\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$  is a orthogonal basis of  $\mathbb{R}^3$  and

$\left\{ \vec{v}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$  is a orthonormal

basis of  $\mathbb{R}^3$

The singular values of  $A$  are  $\sigma_1 = \sqrt{3}$ ,  $\sigma_2 = 1$  and  $\sigma = 0$ . Then define

$$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i,$$

where  $i = 1, 2$ .

Therefore,  $\vec{u}_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} -3 \\ 3 \end{bmatrix}$  and  $\vec{u}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ .

This is clearly a orthonormal basis of  $\mathbb{R}^2$ . Thus the singular value decomposition of  $A$  is given as follows:

$$A = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}^T.$$