# Norm

Tuan Nguyen

Ngày 24 tháng 12 năm 2020

# Overview

# Linear mapping

For vector spaces V, W , a linear mapping $\Theta : V \to W$ is called a linear mapping (or vector space homomorphism/ linear transformation) if:

$$\forall x, y \in V, \forall \alpha, \beta : \Theta(\alpha x + \beta y) = \alpha\Theta(x) + \beta\Theta(y)$$

The mapping $\Theta : \mathbb{R}^2 \to \mathbb{C}, \Theta(x) = x_1 + ix_2$ is a homomorphism.

# Linear mapping (cont.)

Consider a linear mapping $\Theta : V \to W$, where V, W can be arbitrary sets. Then $\Theta$ is called:

- Injective if $x, y \in V : \Theta(x) = \Theta(y) \Rightarrow x = y$.
- Surjective if $\Theta(\mathrm{V}) = \mathrm{W}$.
- Bijective if it is injective and surjective.

Special cases of linear mappings between vector spaces V and W :

- Isomorphism: $\Theta : V \to W$ linear and bijective
- Endomorphism: $\Theta : V \to V$ linear
- Automorphism: $\Theta : V \to V$ linear and bijective

# Coordinates

Consider a vector space V and an ordered basis B $= (b_1, ..., b_n)$ of V . For any $x \in V$ we obtain a unique representation (linear combination)

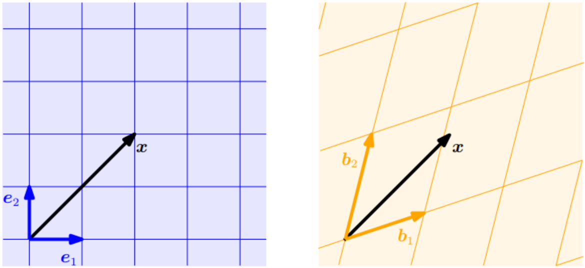
$$x = \alpha_1 v_1 + ... + \alpha_n v_n$$

of x with respect to B. Then $\alpha_1, ..., \alpha_n$ are the coordinates of x with respect to B, and the vector

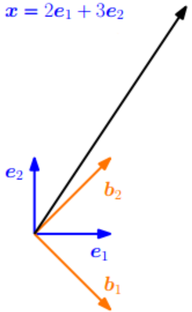
$$\alpha = \begin{bmatrix} \alpha_1 \\ ... \\ \alpha_n \end{bmatrix}$$

is the coordinate vector/coordinate representation of x with respect to the ordered basis B.

# Coordinates (cont.)




Hình 1: Different coordinate representations of a vector x




Hình 2: Different coordinate representations of a vector x

What is the coordinate of x in base $(b_1, b_2)$?

# Transformation matrix

Consider vector spaces V; W with corresponding (ordered) bases B $= (b_1, ..., b_n)$ and C $= (c_1, ..., c_m)$. Moreover, we consider a linear mapping $\Theta : V \to W$ . For $j \in \{1, ..., n\}$:

$$\Theta(b_j) = \sum_{i=1}^{m} \alpha_{ij} c_j$$

is the unique representation of $\Theta(b_j)$ with respect to C. Then, we call the matrix $A_\theta \in \mathbb{R}^{m \times n}$, whose elements are given by:

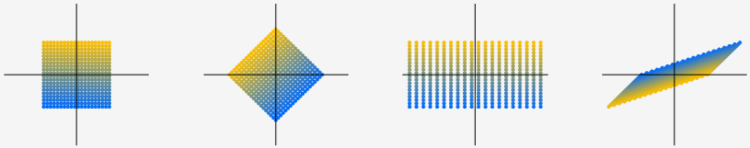
$$A_\theta(i, j) = \alpha_{ij}$$

the transformation matrix of $\Theta$ (with respect to the ordered bases B of V and and C of W).

If $\hat{x}$ is the coordinate vector of $x \in V$ with respect to B and $\hat{y}$ the coordinate vector of $y = \Theta(x) \in W$ with respect to C, then:

$$\hat{y} = A_\theta \hat{x}$$

# Transformation matrix (cont.)



(a) Original data. (b) Rotation by 45°. (c) Stretch along the horizontal axis. (d) General linear mapping.

We consider three linear transformations of a set of vectors in $\mathbb{R}^2$ with the transformation matrices

$$\boldsymbol{A}_1 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{A}_3 = \frac{1}{2}\begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix}.$$

Hình 3: Linear transformation of vector

# Basis change

For a linear mapping $\Theta : V \to W$, ordered bases: B $= (b_1, ..., b_n)$; $\tilde{B} = (\tilde{b_1}, ..., \tilde{b_n})$ of V and C $= (c_1, ..., c_m)$; $\tilde{C} = (\tilde{c_1}, ..., \tilde{c_m})$ of W , and a transformation matrix $A_\Theta$ of $\Theta$ with respect to B and C, the corresponding transformation matrix $\tilde{A_\Theta}$ with respect to the bases $\tilde{B}$ and $\tilde{C}$ is given as:

$$\tilde{A_\Theta} = T^{-1} A_\Theta S$$

Here, $S \in \mathbf{R}^{n \times n}$ is the transformation matrix of $id_V$ that maps coordinates with respect to $\tilde{B}$ onto coordinates with respect to B, and $T \in \mathbf{R}^{m \times m}$ is the transformation matrix of $id_W$ that maps coordinates with respect to $\tilde{C}$ onto coordinates with respect to C.

# Basis change (cont.)

$$
\begin{tikzcd}
\text{Vector spaces} & V \arrow[rr, "\Phi"] & & W & & V \arrow[rr, "\Phi"] & & W \\
& B \arrow[rr, "\Phi_{CB}", "\boldsymbol{A}_{\Phi}"'] & & C & & B \arrow[rr, "\Phi_{CB}", "\boldsymbol{A}_{\Phi}"'] & & C \arrow[d, "\boldsymbol{T}^{-1}"', "\Xi_{\tilde{C}C} = \Xi^{-1}_{C\tilde{C}}"] \\
\text{Ordered bases} & \tilde{B} \arrow[u, "\Psi_{B\tilde{B}}", "\boldsymbol{S}"'] \arrow[rr, "\tilde{\boldsymbol{A}}_{\Phi}", "\Phi_{\tilde{C}\tilde{B}}"'] & & \tilde{C} \arrow[u, "\boldsymbol{T}", "\Xi_{C\tilde{C}}"'] & & \tilde{B} \arrow[u, "\Psi_{B\tilde{B}}", "\boldsymbol{S}"'] \arrow[rr, "\tilde{\boldsymbol{A}}_{\Phi}", "\Phi_{\tilde{C}\tilde{B}}"'] & & \tilde{C}
\end{tikzcd}
$$

Hình 4: Change basis

# Basis change (cont.)

Consider a linear mapping $\Theta : \mathbb{R}^3 \to \mathbb{R}^4$ whose transformation matrix is

$$A_\Theta = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}$$

with respect to the standard bases

$$B = (\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}), C = (\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix})$$

$$\tilde{B} = (\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}), \tilde{C} = (\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix})$$

# Image and Kernel

For $\Theta : V \to W$ , we define the kernel/null space

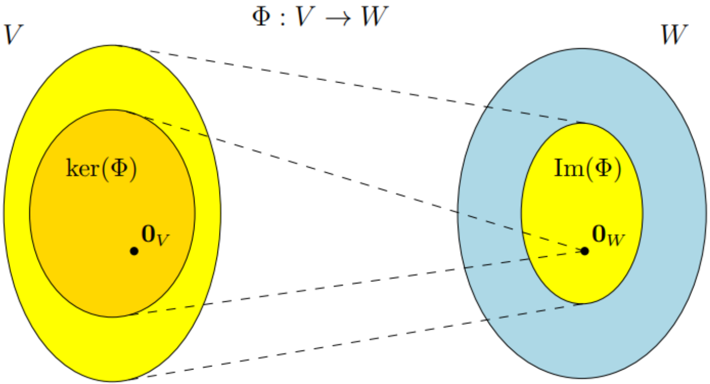
$$\ker(\Theta) := \Theta^{-1}(0_W) = \{v \in V : \Theta(v) = 0_W\}$$

and the image/range

$$\mathrm{Im}(\Theta) := \Theta(V) = \{w \in W | \exists v \in V : \Theta(v) = w\}$$

We also call V and W also the domain and codomain of $\Theta$ respectively.

- It always holds that $\Theta(0_V) = 0_W$ and, therefore, $0_V \in \ker(\Theta)$. In particular, the null space is never empty.
- $\mathrm{Im}(\Theta) \subseteq W$ is a subspace of W , and $\ker(\Theta) \subseteq V$ is a subspace of V .




Hình 5: Illustration of image and kernel

# Image and Kernel (cont.)

Null Space and Column Space: Let us consider $A \in \mathbb{R}^{m \times n}$ and a linear mapping $\Theta : \mathbb{R}^m \to \mathbb{R}^n, x \to Ax$.

- For A $= [a_1, ..., a_n]$, where $a_i$ are the columns of A, we obtain Im($\Theta$) $= \{Ax : x \in \mathbb{R}^n\} = \{\sum_{i=1}^n x_i a_i : x_1, ..., x_n \in \mathbb{R}^n\} = span[a_1, ..., a_n] \subseteq \mathbb{R}^m$
- rk(A) = dim(Im($\Theta$))
- The kernel/null space ker($\Theta$) is the general solution to the homogeneous system of linear equations Ax = 0 and captures all possible linear combinations of the elements in $\mathbb{R}^n$ that produce $0 \in \mathbb{R}^m$.

The mapping:

$$\Theta : \mathbb{R}^4 \to \mathbb{R}^2, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix}$$

Determine Im($\Theta$), Ker($\Theta$)

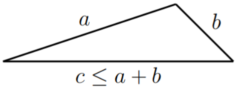
# Norm

A norm on a vector space V is a function

$$\|\cdot\| : V \to R, x \to \|x\|$$

which assigns each vector x its length $\|x\|$, such that for all $\lambda \in \mathbb{R}$ and $x, y \in \mathbb{R}$ the following hold:

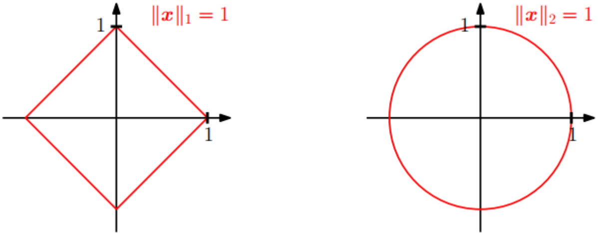
- $\|\lambda x\| = |\lambda| \|x\|$
- $\|x + y\| \leqslant \|x\| + \|y\|$
- $\|x\| \geqslant 0, \|x\| = 0 \Leftrightarrow x = 0$




Hình 6: Triangle inequality

For $x \in \mathbb{R}^n$, we define:

- Manhattan Norm ($l_1$): $\|x\|_1 := \sum_{i=1}^{n} |x_i|$
- Euclidean Norm ($l_2$): $\|x\|_2 := \sqrt{\sum_{i=1}^{n} x_i^2}$




Hình 7: Example of $l_1, l_2$ norm

# Dot product

For $x, y \in \mathbb{R}^n$, the scalar product/dot product of two vectors x, y:

$$x^T y = \sum_{i=1}^{n} x_i y_i$$

# Inner product

A bilinear mapping $\Omega$ is a mapping with two arguments, and it is linear in each argument, i.e., when we look at a vector space V then it holds that for all $x, y, z \in V, \lambda, \psi \in \mathbb{R}$ that:

$$\Omega(\lambda x + \psi y, z) = \lambda\Omega(x, z) + \psi\Omega(y, z)$$
$$\Omega(x, \lambda y + \psi z) = \lambda\Omega(x, y) + \psi\Omega(x, z)$$

Let V be a vector space and bilinear mapping $\Omega : V \times V \to \mathbb{R}$:

- $\Omega$ is symmetric if $\Omega(x, y) = \Omega(y, x)$ for all $x, y \in V$
- $\Omega$ is positive definite if

$$\forall x \in V/\{0\} : \Omega(x, x) > 0, \Omega(0, 0) = 0$$

# Inner product (cont.)

Let V be a vector space and $\Omega : V \times V \to \mathbb{R}$ be a bilinear mapping that takes two vectors and maps them onto a real number. Then:

- A positive definite, symmetric bilinear mapping $\Omega : V \times V \to \mathbb{R}$ is called an inner product on V. We typically write $\langle x, y \rangle$ instead of $\Omega(x, y)$.

Dot product is the most common inner product, but inner product probably is not dot product. For example:

$$\langle x, y \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2$$

# Symmetric, Positive Definite Matrix

A symmetric matrix $A \in \mathbb{R}^{m \times n}$ that satisfies

$$\forall x \in V/\{0\} : x^T Ax > 0(1)$$

is called symmetric, positive definite, or definite just positive definite. If only $\geqslant$ holds in (1), then A is called symmetric, positive semidefinite.

Are the following matrix positive definite:

$$A_1 = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix}, A_2 = \begin{bmatrix} 9 & 6 \\ 6 & 3 \end{bmatrix}$$

Properties:

- Kernel (null space) of A consists only 0.
- The diagonal elements $a_{ii}$ of A are positive because $a_{ii} = e_i^T Ae_i$

# Length

The length of vector x is defined as:

$$\|x\| := \sqrt{\langle x, x \rangle}$$

For example, let us take x $= [1,\ 1]^T$

If we use dot product as inner product $\|x\| = \sqrt{1^2 + 1^2} = \sqrt{2}$ Let choose different inner product

$\langle x, y \rangle = x^T \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} y = x_1 y_1 - \frac{1}{2}(x_1 y_2 + x_2 y_1) + x_2 y_2$

$\Rightarrow \|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 - x_1 x_2 + x_2^2} = 1$

$\Rightarrow$x is shorter in this inner product than with dot product.

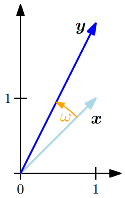
# Angle

$\omega$ is the angle between two vectors x, y is defined as:

$$cos(\omega) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

For example, $x = [1, 1]^T, y = [1, 2]^T \in \mathbb{R}^2$ and we ues dot product as inner product:

$$cos(\omega) = \frac{x^T y}{\|x\| \|y\|} = \frac{3}{\sqrt{10}}$$



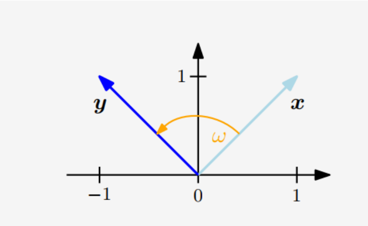

Hình 8: Angle $\omega$ between vectors x and y

# Orthogonality

Two vectors x and y are orthogonal if and only if $\langle x, y\rangle = 0$, and we write $x \perp y$. If additionally $\|x\| = \|y\| = 1$, i.e., the vectors are unit vectors, then x and y are orthonormal.

For example, $x = [1, 1]^T, y = [-1, 1]^T \in \mathbb{R}^2$, if we use dot product as inner product:

$$\langle x, y\rangle = x^T y = 0 \Rightarrow cos(\omega) = 0 \Rightarrow \omega = 90^o$$




Hình 9: orthogonal vector

# Orthogonal matrix

A square matrix $A \in \mathbb{R}^{n \times n}$ is an orthogonal matrix if and only if its columns (rows) are orthonormal so that

$$AA^T = I_n = A^T A$$

which implies that $A^T = A^{-1}$

Remark:

- The length of a vector x is not changed when transforming it using an orthogonal matrix A:

$$\|Ax\|^2 = (Ax)^T(Ax) = x^T A^T Ax = x^T x = \|x\|^2$$

- The angle betweet vector x and y is not changed when transforming them using an orthogonal matrix A:

$$\cos(\omega) = \frac{\langle Ax, Ay \rangle}{\|Ax\|\|Ay\|} = \frac{(Ax)^T(Ax)}{\sqrt{(Ax)^T(Ax)(Ay)^T(Ay)}} = \frac{x^T y}{\|x\|\|y\|}$$

It means orthogonal matrix A preserves both angles and distances

# Orthonormal Basis

Consider an n-dimensional vector space V and a basis $b_1, ..., b_n$ of V .
If

$$\langle b_i, b_j \rangle = 0 \text{ if } i \neq j \qquad (1)$$
$$\langle b_i, b_i \rangle = 1$$

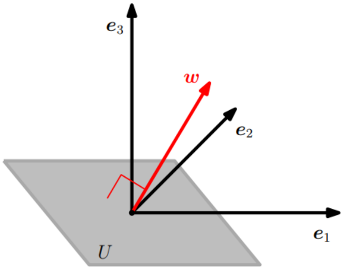
for all i, j = 1,...,n then the basis is called an orthonormal basis (ONB). If only (1) is satisfied, then the basis is called an orthogonal basis.

In $\mathbb{R}^2$, the vectors:

$$b_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, b_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Are they orthogonal/orthogomal basis?

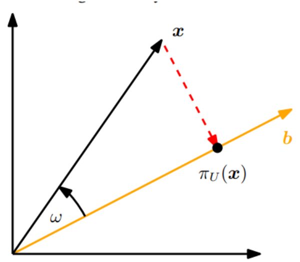
Consider a D-dimensional vector space V and an M-dimensional subspace U $\subseteq$ V. Then its orthogonal complement $U^{\perp}$ is a (D - M)-dimensional orthogonal subspace of V and contains all vectors in V that are orthogonal to every complement vector in U, $U \cap U^{\perp} = \{0\}$.




Hình 10: Orthogonal complement of two dimensional space

Let V be a vector space and U $\subseteq$ V a subspace of V . A linear mapping $\pi : V \to U$ is called a projection if $\pi^2 = \pi \circ \pi = \pi$

**Projection onto one-dimensional subspaces (lines)** The line is a one-dimensional subspace $U \subseteq \mathbb{R}^n$ spanned by b. When we project $x \in \mathbb{R}^n$ onto U, we seek the vector $\pi_U(x)$ that is closest to x.




(a) Projection of $\boldsymbol{x} \in \mathbb{R}^2$ onto a subspace $U$ with basis vector $\boldsymbol{b}$.

# Projection (cont.)

Properties of the projection $\pi_U(x)$

- The projection $\pi_U(x)$ is closest to x, where "closest" implies that the distance $\|x - \pi_U(x)\|$ is minimal $\Rightarrow \pi_U(x) - x$ from $\pi_U(x)$ to x is orthogonal to U $\Rightarrow \langle \pi_U(x) - x, b \rangle = 0$
- $\pi_U(x) \in U \Rightarrow \pi_U(x) = \lambda b$

**We find $\lambda$**

$$\langle x - \pi_U(x), b \rangle = 0 \Leftrightarrow \langle x - \lambda b, b \rangle = 0 \Leftrightarrow \langle x, b \rangle = \lambda \langle b, b \rangle$$

$$\Leftrightarrow \lambda = \frac{\langle x, b \rangle}{\langle b, b \rangle} = \frac{\langle b, x \rangle}{\|b\|^2}$$

If we choose inner product as dot product

$$\lambda = \frac{\langle b, x \rangle}{\|b\|^2} = \frac{b^T x}{\|b\|^2} \Rightarrow \pi_U(x) = \frac{b^T x}{\|b\|^2} b$$

If $\|b\| = 1 \Rightarrow \lambda = b^T x$

# Projection (cont.)

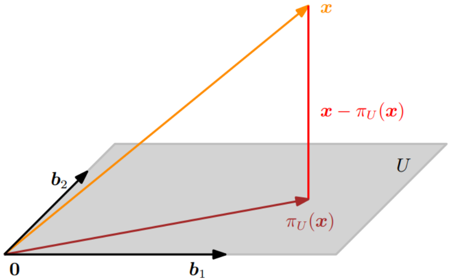
**We find transforation matrix $P_\pi$**

$$\pi_U(x) = \frac{b^T x}{\|b\|^2} b = \frac{bb^T}{\|b\|^2} x$$
$$\Rightarrow P_\pi = \frac{bb^T}{\|b\|^2}$$

The projection matrix $P_\pi$ projects any vector $x \in \mathbb{R}^n$ onto the line through the origin with direction b

### Projection onto general subspaces

we look at orthogonal projections of vectors $x \in \mathbb{R}^n$ onto lower-dimensional subspaces $U \subseteq \mathbb{R}^n$ with dim(U) = m > 1




Hình 11: Projection onto 2d space

# Projection (cont.)

Assume that $(b_1, \dots, b_m)$ is an ordered basis of U. We have:

$$\pi_U(x) = \sum_{i=1}^{m} \lambda_i b_i = B\lambda$$

where $B = [b_1, \dots, b_m] \in \mathbb{R}^{n \times m}, \lambda = [\lambda_1, \dots, \lambda_m]^T \in \mathbb{R}^m$, $x - \pi_U(x)$ is orthogonal to U $\Rightarrow x - \pi_U(x)$ is orthogonal to $b_1, \dots, b_m$

$$\langle b_1, x - \pi_U(x) \rangle = 0$$
$$\dots$$
$$\langle b_m, x - \pi_U(x) \rangle = 0$$

# Projection (cont.)

With $\pi_U(x) = B\lambda$, we can write as:

$$\langle b_1, x - B\lambda \rangle = 0$$
$$...$$
$$\langle b_m, x - B\lambda \rangle = 0$$

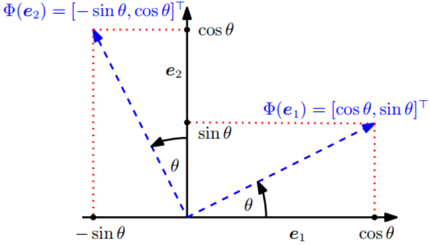
$$\Rightarrow \begin{bmatrix} b_1^T \\ ... \\ b_n^T \end{bmatrix} \begin{bmatrix} x - B\lambda \end{bmatrix} = 0 \Leftrightarrow B^T(x - B\lambda) = 0 \Leftrightarrow B^T B\lambda = B^T x$$

Since $b_1, ..., b_m$ are a basis of U and, therefore, linearly independent => $B^T B \in \mathbb{R}^{m \times m}$ is regular and can be inverted. This allows us to solve for the coefficients:

$$\lambda = (B^T B)^{-1} B^T x \Rightarrow \pi_U(x) = B(B^T B)^{-1} B^T x$$

A rotation is a linear mapping that rotates a plane by an angle $\theta$ about the origin, For a positive angle $\theta > 0$, we rotate in a counterclockwise direction. Transformation matrix:

$$R(\theta) = [\Theta(e_1), \Theta(e_2)] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$




Hình 12: Rotation in $\mathbb{R}^2$