شبکه‌های عصبی مصنوعی

درس ۶

تبدیل‌های خطی برای شبکه‌های عصبی

Linear Transformations for Neural Networks

کاظم فولادی قلعه

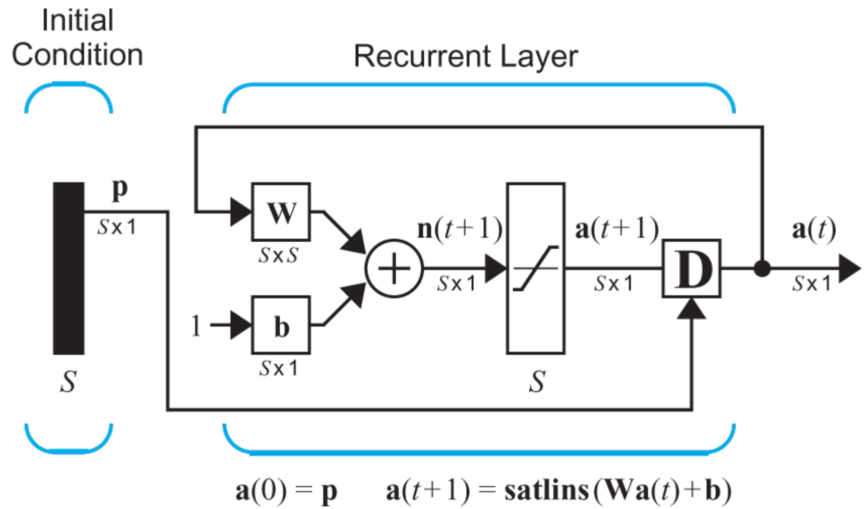
دانشکده مهندسی، پردیس فارابی

دانشگاه تهران

http://courses.fouladi.ir/nn

Linear Transformations

Hopfield Network Questions

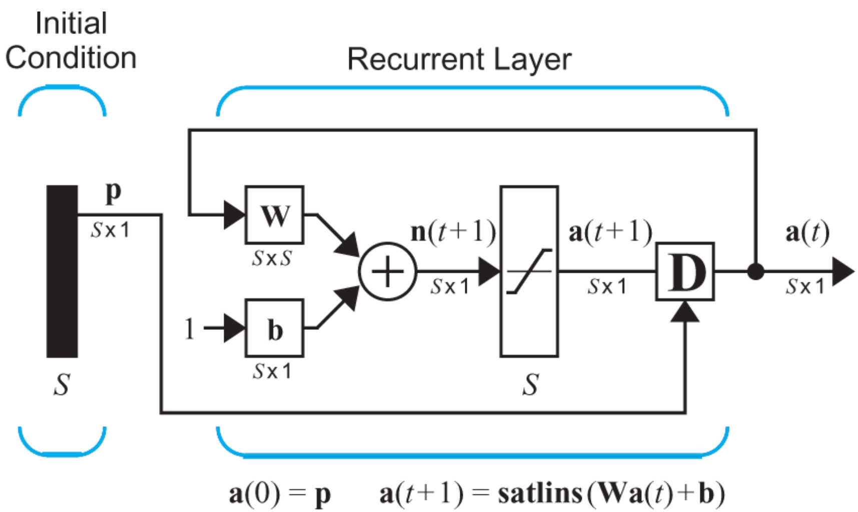


- The network output is repeatedly multiplied by the weight matrix W.
- What is the effect of this repeated operation?
- Will the output converge, go to infinity, oscillate?
- In this chapter we want to investigate matrix multiplication, which represents a general linear transformation.

پرسش‌هایی درباره‌ی شبکه‌ی هاپفیلد

ضرورت مطالعه‌ی تبدیل‌های خطی

HOPFIELD NETWORK QUESTIONS



- خروجی شبکه به‌طور مکرر در ماتریس وزن **W** ضرب می‌شود.
- اثر این عمل تکراری چیست؟
- آیا خروجی همگرا می‌شود؟ به سمت بی‌نهایت می‌رود؟ یا نوسان می‌کند؟
- ضرب یک ماتریس وزن **W** در بردار ورودی **p** یکی از اعمال کلیدی است که توسط شبکه‌های عصبی انجام می‌شود.
- این ضرب نمونه‌ای از یک تبدیل خطی است ⇐ نیاز به بررسی تبدیل خطی عمومی و تعیین مشخصات بنیادی آن.

تبدیل‌های خطی برای شبکه‌های عصبی

۱

تبدیل‌های خطی

6

Linear Transformations

A **transformation** consists of three parts:

1. A set of elements $X = \{x_i\}$, called the domain,
2. A set of elements $Y = \{y_i\}$, called the range, and
3. A rule relating each $x_i \in X$ to an element $y_i \in Y$.

A transformation is **linear** if:

1. For all $x_1, x_2 \in X$, $A(x_1 + x_2) = A(x_1) + A(x_2)$,
2. For all $x \in X$, $a \in \Re$, $A(a\,x) = a\,A(x)$.

تبدیل‌های خطی

LINEAR TRANSFORMATIONS

یک **تبدیل** سه جزء دارد:

A **transformation** consists of three parts:

دامنه — 1. A set of elements $X = \{x_i\}$, called the domain,

برد — 2. A set of elements $Y = \{y_i\}$, called the range, and

ضابطه — 3. A rule relating each $x_i \in X$ to an element $y_i \in Y$.

خواص یک **تبدیل** خطی:

A transformation is **linear** if:

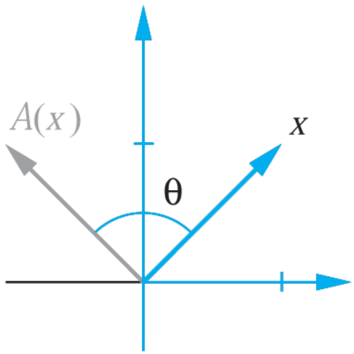
۱) جمع‌پذیری — 1. For all $x_1, x_2 \in X$, $A(x_1 + x_2) = A(x_1) + A(x_2)$,

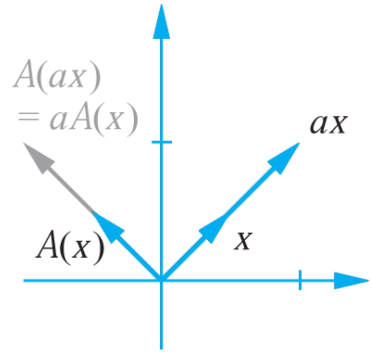
۲) همگنی — 2. For all $x \in X$, $a \in \Re$, $A(a\,x) = a\,A(x)$.

Example - Rotation

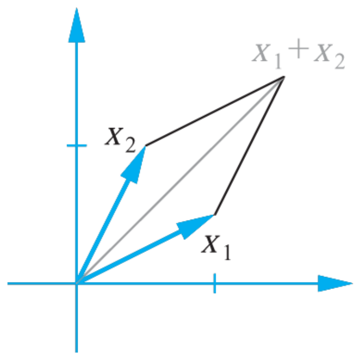
Is rotation linear?

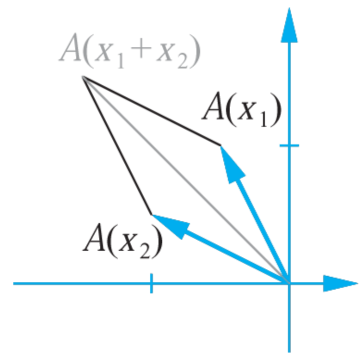
1.





2.



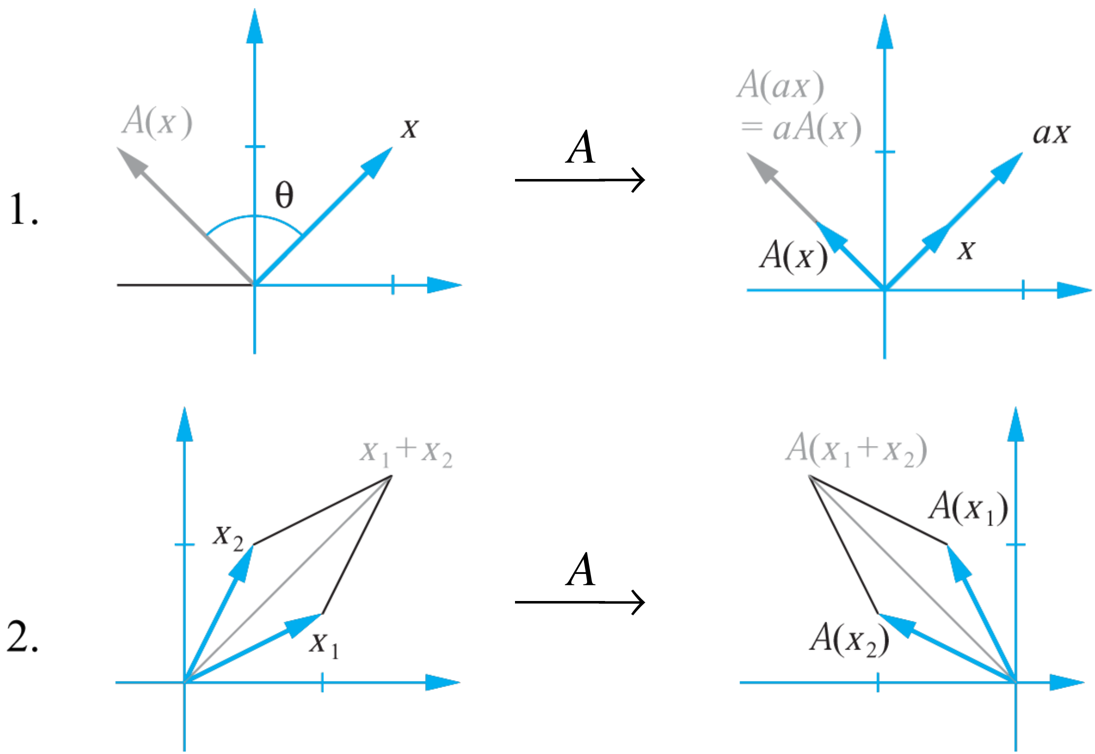


تبدیل‌های خطی

مثال : تبدیل دوران

Linear Transformations

تبدیل دوران به‌اندازه‌ی زاویه‌ی θ در $\mathbb{R}^2$ یک تبدیل خطی است.



۲

بازنمایی‌های ماتریسی برای تبدیل‌های خطی

Matrix Representation - (1)

6

Any linear transformation between two finite-dimensional vector spaces can be represented by matrix multiplication.

Let $\{v_1, v_2, \ldots, v_n\}$ be a basis for X, and let $\{u_1, u_2, \ldots, u_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x_i v_i \qquad y = \sum_{i=1}^{m} y_i u_i$$

Let $A{:}X \rightarrow Y$

$$A(x) = y$$

$$A\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{i=1}^{m} y_i u_i$$

بازنمایی ماتریسی

(۱ از ۳)

Matrix Representation

هر تبدیل خطی بین دو فضای برداری متناهی–ابعاد می‌تواند توسط ضرب ماتریسی بازنمایی شود.

Let $\{v_1, v_2, ..., v_n\}$ be a basis for X, and let $\{u_1, u_2, ..., u_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x_i v_i \qquad y = \sum_{i=1}^{m} y_i u_i$$

Let $A: X \rightarrow Y$

$$A(x) = y$$

$$A\left(\sum_{j=1}^{n} x_j v_j\right) = \sum_{i=1}^{m} y_i u_i$$

Matrix Representation - (2)

Since A is a linear operator,

$$\sum_{j=1}^{n} x_j A(v_j) = \sum_{i=1}^{m} y_i u_i$$

Since the u_i are a basis for Y,

$$A(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

(The coefficients a_{ij} will make up the matrix representation of the transformation.)

$$\sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} u_i = \sum_{i=1}^{m} y_i u_i$$

بازنمایی ماتریسی

(۲ از ۳)

Matrix Representation

Since A is a linear operator,

$$\sum_{j=1}^{n} x_j A(v_j) = \sum_{i=1}^{m} y_i u_i$$

Since the u_i are a basis for Y,

$$A(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

(The coefficients a_{ij} will make up the matrix representation of the transformation.)

$$\sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} u_i = \sum_{i=1}^{m} y_i u_i$$

6 Matrix Representation - (3)

$$\sum_{i=1}^{m} u_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{m} y_i u_i$$

$$\sum_{i=1}^{m} u_i \left(\sum_{j=1}^{n} a_{ij} x_j - y_i \right) = 0$$

Because the u_i are independent,

$$\sum_{j=1}^{n} a_{ij} x_j = y_i$$

This is equivalent to matrix multiplication.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

بازنمایی ماتریسی

(۳ از ۳)

Matrix Representation

$$\sum_{i=1}^{m} u_i \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{m} y_i u_i$$

$$\sum_{i=1}^{m} u_i \left(\sum_{j=1}^{n} a_{ij} x_j - y_i \right) = 0$$

Because the u_i are independent,

$$\sum_{j=1}^{n} a_{ij} x_j = y_i$$

This is equivalent to matrix multiplication.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Summary

- A linear transformation can be represented by matrix multiplication.
- To find the matrix which represents the transformation we must transform each basis vector for the domain and then expand the result in terms of the basis vectors of the range.

$$A(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

Each of these equations gives us one column of the matrix.

بازنمایی ماتریسی

خلاصه

MATRIX REPRESENTATION

هر تبدیل خطی می‌تواند توسط ضرب ماتریسی بازنمایی شود.

برای یافتن ماتریس تبدیل باید هر یک از بردارهای پایه‌ی دامنه را تبدیل کنیم
و سپس حاصل را بر حسب بردارهای پایه‌ی برد بسط بدهیم.

$$\mathrm{A}(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{u}_i$$

هر معادله به فرم فوق، یک ستون از ماتریس را به ما می‌دهد.

نکته: اگر مجموعه‌ی پایه برای دامنه یا برد را عوض کنیم، بازنمایی ماتریسی نیز تغییر می‌کند.

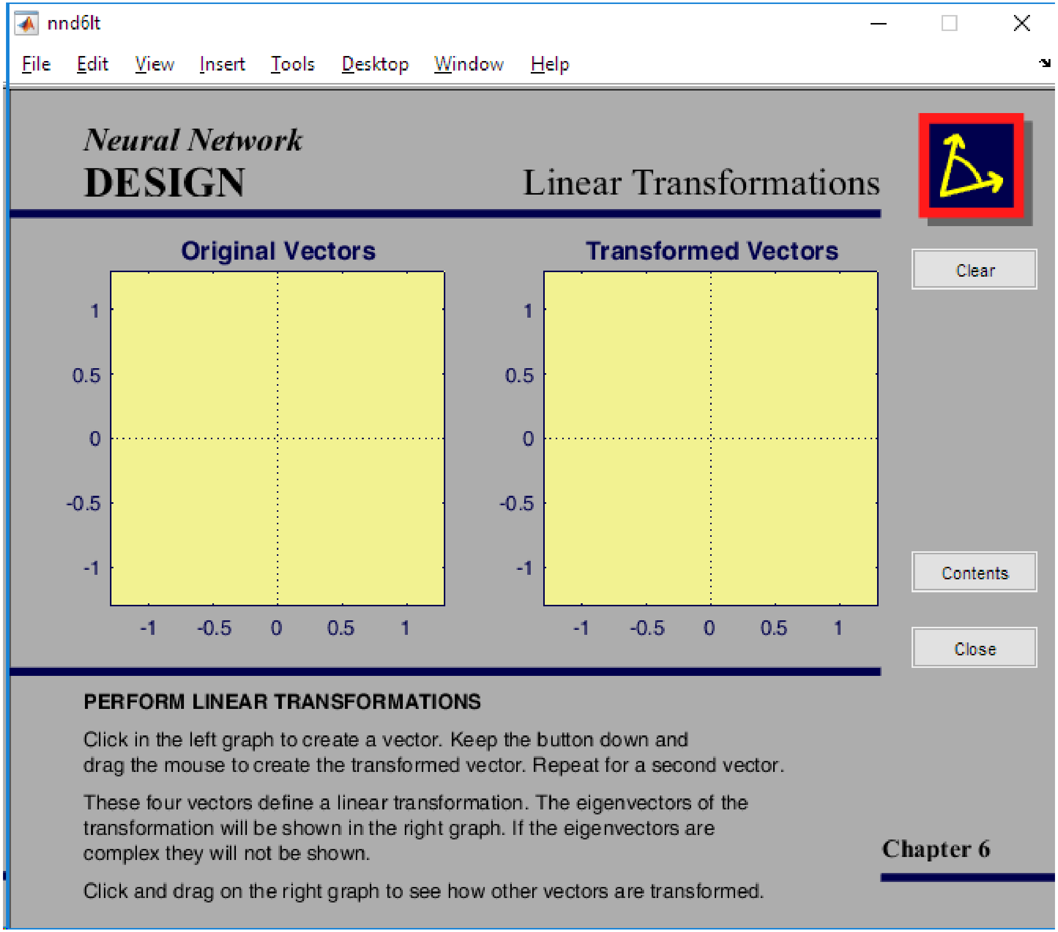
بازنمایی ماتریسی

مثال: دوران

ROTATION

ماتریس تبدیل برای دوران به‌اندازه‌ی θ

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

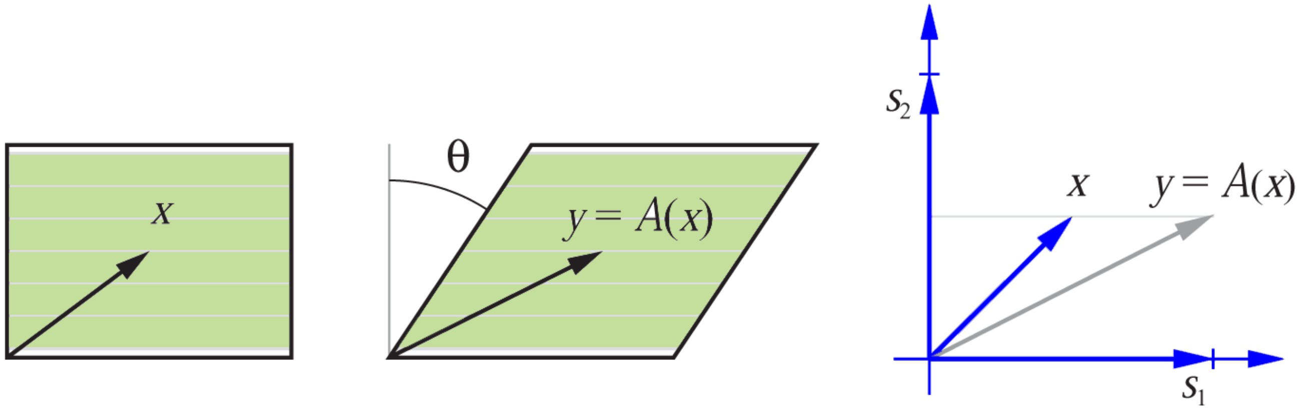


MATLAB

```
>> nnd6lt
```

Example - (1)

Stand a deck of playing cards on edge so that you are looking at the deck sideways. Draw a vector x on the edge of the deck. Now “skew” the deck by an angle θ, as shown below, and note the new vector $y = A(x)$. What is the matrix of this transformation in terms of the standard basis set?

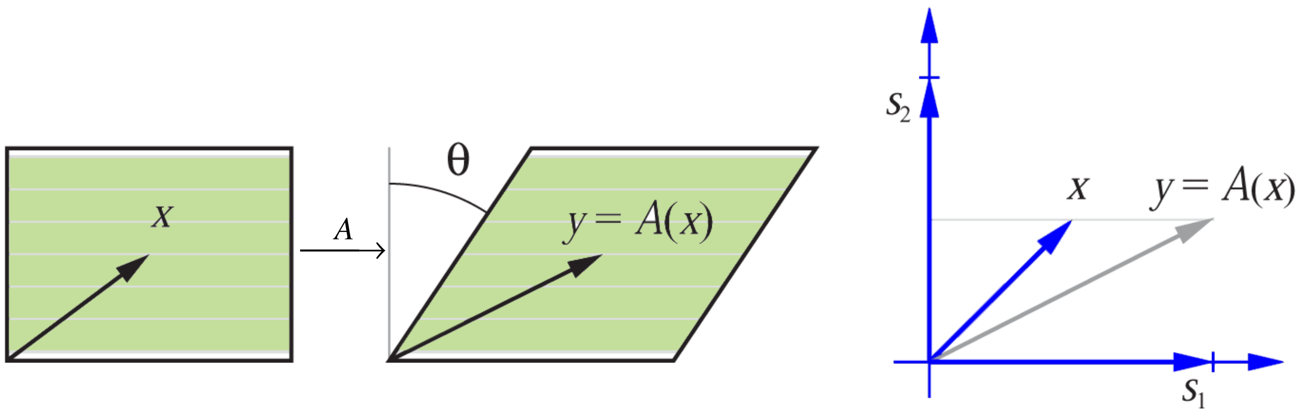


بازنمایی ماتریسی

مثال: مایل‌سازی (۱ از ۵)

SKEW

Stand a deck of playing cards on edge so that you are looking at the deck sideways. Draw a vector x on the edge of the deck. Now "skew" the deck by an angle θ, as shown below, and note the new vector $y = A(x)$. What is the matrix of this transformation in terms of the standard basis set?



Example - (2)

To find the matrix we need to transform each of the basis vectors.

$$A(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

We will use the standard basis vectors for both the domain and the range.

$$A(s_j) = \sum_{i=1}^{2} a_{ij} s_i = a_{1j} s_1 + a_{2j} s_2$$

بازنمایی ماتریسی

مثال: مایل‌سازی (۲ از ۵)

SKEW

To find the matrix we need to transform each of the basis vectors.

$$A(v_j) = \sum_{i=1}^{m} a_{ij} u_i$$

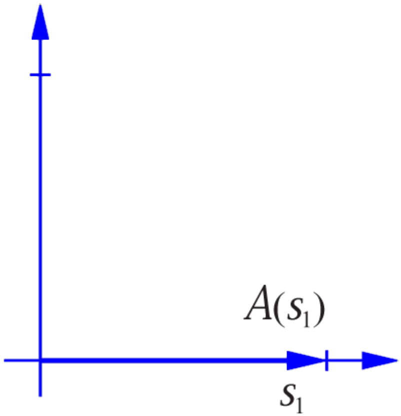
We will use the standard basis vectors for both the domain and the range.

$$A(s_j) = \sum_{i=1}^{2} a_{ij} s_i = a_{1j} s_1 + a_{2j} s_2$$

Example - (3)

We begin with s_1:

If we draw a line on the bottom card and then skew the deck, the line will not change.



$$A(s_1) = 1s_1 + 0s_2 = \sum_{i=1}^{2} a_{i1}s_i = a_{11}s_1 + a_{21}s_2$$

This gives us the first column of the matrix.

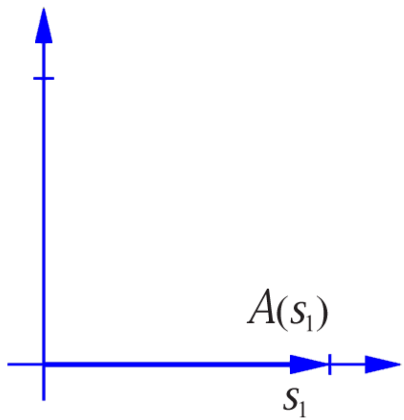
بازنمایی ماتریسی

مثال: مایل‌سازی (۳ از ۵)

SKEW

We begin with s_1:

If we draw a line on the bottom card and then skew the deck, the line will not change.

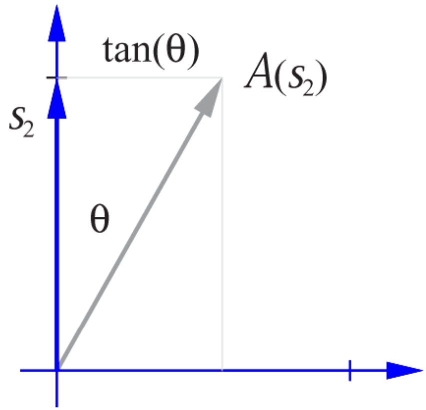


$$A(s_1) = 1s_1 + 0s_2 = \sum_{i=1}^{2} a_{i1}s_i = a_{11}s_1 + a_{21}s_2$$

This gives us the first column of the matrix.

Example - (4)

Next, we skew s_2:



$$A(s_2) = \tan(\theta)s_1 + 1s_2 = \sum_{i=1}^{2} a_{i2}s_i = a_{12}s_1 + a_{22}s_2$$

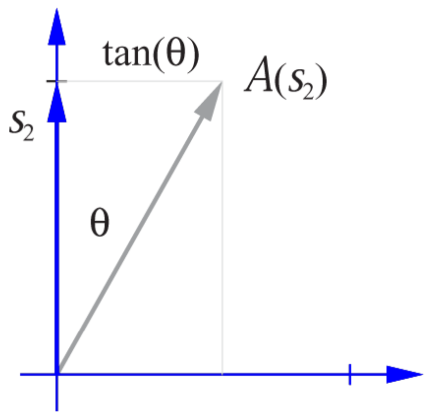
This gives us the second column of the matrix.

بازنمایی ماتریسی

مثال: مایل‌سازی (۴ از ۵)

SKEW

Next, we skew s_2:



$$A(s_2) = \tan(\theta)s_1 + 1s_2 = \sum_{i=1}^{2} a_{i2}s_i = a_{12}s_1 + a_{22}s_2$$

This gives us the second column of the matrix.

Example - (5)

The matrix of the transformation is:

$$\mathbf{A} = \begin{bmatrix} 1 & \tan(\theta) \\ 0 & 1 \end{bmatrix}$$

بازنمایی ماتریسی

مثال: مایل‌سازی (۵ از ۵)

SKEW

The matrix of the transformation is:

$$\mathbf{A} = \begin{bmatrix} 1 & \tan(\theta) \\ 0 & 1 \end{bmatrix}$$

تبدیل‌های خطی برای شبکه‌های عصبی

۳

تغییر پایه‌ها

Change of Basis

Consider the linear transformation $A{:}X \rightarrow Y$. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis for X, and let $\{u_1, u_2, \ldots, u_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x_i v_i \qquad\qquad y = \sum_{i=1}^{m} y_i u_i$$

$$A(x) = y$$

The matrix representation is:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

تغییر پایه‌ها

Change of Basis

Consider the linear transformation $A{:}X \rightarrow Y$. Let $\{v_1, v_2, ..., v_n\}$ be a basis for X, and let $\{u_1, u_2, ..., u_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x_i v_i \qquad y = \sum_{i=1}^{m} y_i u_i$$

$$A(x) = y$$

The matrix representation is:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

$$\mathbf{Ax} = \mathbf{y}$$

New Basis Sets

Now let's consider different basis sets. Let $\{t_1, t_2, ..., t_n\}$ be a basis for X, and let $\{w_1, w_2, ..., w_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x'_i t_i \qquad y = \sum_{i=1}^{m} y'_i w_i$$

The new matrix representation is:

$$\begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{bmatrix}$$

$$\mathbf{A'x'} = \mathbf{y'}$$

تغییر پایه‌ها

مجموعه‌های پایه‌ی جدید

New Basis Sets

Now let’s consider different basis sets. Let $\{t_1, t_2, \ldots, t_n\}$ be a basis for X, and let $\{w_1, w_2, \ldots, w_m\}$ be a basis for Y.

$$x = \sum_{i=1}^{n} x'_i t_i \qquad\qquad y = \sum_{i=1}^{m} y'_i w_i$$

The new matrix representation is:

$$\begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ a'_{21} & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ a'_{m1} & a'_{m2} & \cdots & a'_{mn} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{bmatrix}$$

$$\mathbf{A'x'} = \mathbf{y'}$$

6

How are A and A' related?

Expand t_i in terms of the original basis vectors for X.

$$t_i = \sum_{j=1}^{n} t_{ji} v_j \qquad \mathbf{t}_i = \begin{bmatrix} t_{1i} \\ t_{2i} \\ \vdots \\ t_{ni} \end{bmatrix}$$

Expand w_i in terms of the original basis vectors for Y.

$$w_i = \sum_{j=1}^{m} w_{ji} u_j \qquad \mathbf{w}_i = \begin{bmatrix} w_{1i} \\ w_{2i} \\ \vdots \\ w_{mi} \end{bmatrix}$$

تغییر پایه‌ها

ارتباط ماتریس‌های تبدیل A و A’ (۱ از ۲)

New Basis Sets

Expand t_i in terms of the original basis vectors for X.

$$t_i = \sum_{j=1}^{n} t_{ji} v_j \qquad \mathbf{t}_i = \begin{bmatrix} t_{1i} \\ t_{2i} \\ \vdots \\ t_{ni} \end{bmatrix}$$

Expand w_i in terms of the original basis vectors for Y.

$$w_i = \sum_{j=1}^{m} w_{ji} u_j \qquad \mathbf{w}_i = \begin{bmatrix} w_{1i} \\ w_{2i} \\ \vdots \\ w_{mi} \end{bmatrix}$$

6

How are A and A' related?

$$\mathbf{B}_t = \begin{bmatrix}\mathbf{t}_1 & \mathbf{t}_2 & \dots & \mathbf{t}_n\end{bmatrix} \qquad \mathbf{x} = x'_1\mathbf{t}_1 + x'_2\mathbf{t}_2 + \dots + x'_n\mathbf{t}_n = \mathbf{B}_t\mathbf{x}'$$

$$\mathbf{B}_w = \begin{bmatrix}\mathbf{w}_1 & \mathbf{w}_2 & \dots & \mathbf{w}_m\end{bmatrix} \qquad \mathbf{y} = \mathbf{B}_w\mathbf{y}'$$

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad \Longrightarrow \quad \mathbf{A}\mathbf{B}_t\mathbf{x}' = \mathbf{B}_w\mathbf{y}'$$

$$[\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t]\mathbf{x}' = \mathbf{y}'$$

$$\mathbf{A}'\mathbf{x}' = \mathbf{y}' \quad \Longrightarrow \quad \boxed{\mathbf{A}' = [\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t]}$$

Similarity Transform

تغییر پایه‌ها

ارتباط ماتریس‌های تبدیل A و A’ (۲ از ۲)

NEW BASIS SETS

بردار پایه‌ی متقابل

$$\mathbf{B}_t = \begin{bmatrix}\mathbf{t}_1 & \mathbf{t}_2 & \ldots & \mathbf{t}_n\end{bmatrix} \qquad \mathbf{x} = x'_1\mathbf{t}_1 + x'_2\mathbf{t}_2 + \cdots + x'_n\mathbf{t}_n = \mathbf{B}_t\mathbf{x}'$$

$$\mathbf{B}_w = \begin{bmatrix}\mathbf{w}_1 & \mathbf{w}_2 & \ldots & \mathbf{w}_m\end{bmatrix} \qquad \mathbf{y} = \mathbf{B}_w\mathbf{y}'$$

هر ستون یک $\mathbf{w}_i$

$$\mathbf{A}\mathbf{x} = \mathbf{y} \Rightarrow \mathbf{A}\mathbf{B}_t\mathbf{x}' = \mathbf{B}_w\mathbf{y}'$$

$$[\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t]\mathbf{x}' = \mathbf{y}'$$

$$\mathbf{A}'\mathbf{x}' = \mathbf{y}'$$

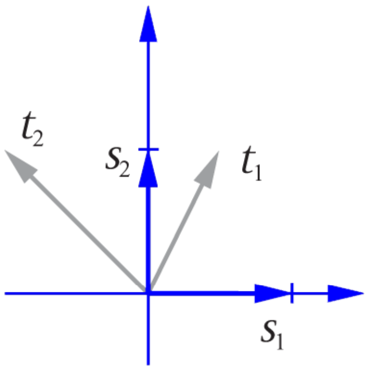
$$\Rightarrow \quad \mathbf{A}' = [\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t]$$

Similarity Transform

تبدیل شباهت

Example - (1)

Take the skewing problem described previously, and find the new matrix representation using the basis set $\{s_1, s_2\}$.



$$t_1 = 0.5s_1 + s_2$$

$$t_2 = -s_1 + s_2$$

$$\mathbf{t}_1 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \qquad \mathbf{t}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad \mathbf{B}_t = \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 \end{bmatrix} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix} \qquad \mathbf{B}_w = \mathbf{B}_t = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

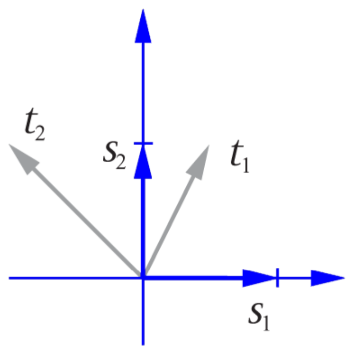
(Same basis for domain and range.)

تغییر پایه‌ها

مثال (۱ از ۳)

NEW BASIS SETS

Take the skewing problem described previously, and find the new matrix representation using the basis set $\{s_1, s_2\}$.



$$t_1 = 0.5 s_1 + s_2$$

$$t_2 = -s_1 + s_2$$

$$\mathbf{t}_1 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \qquad \mathbf{t}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow \mathbf{B}_t = \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 \end{bmatrix} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix} \qquad \mathbf{B}_w = \mathbf{B}_t = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

(Same basis for domain and range.)

Example - (2)

$$\mathbf{A}' = [\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t] = \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & \tan\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} (2/3)\tan\theta + 1 & (2/3)\tan\theta \\ (-2/3)\tan\theta & (-2/3)\tan\theta + 1 \end{bmatrix}$$

For $\theta = 45°$:

$$\mathbf{A}' = \begin{bmatrix} 5/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

تغییر پایه‌ها

مثال (۲ از ۳)

NEW BASIS SETS

$$\mathbf{A}' = [\mathbf{B}_w^{-1}\mathbf{A}\mathbf{B}_t] = \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & \tan\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} (2/3)\tan\theta + 1 & (2/3)\tan\theta \\ (-2/3)\tan\theta & (-2/3)\tan\theta + 1 \end{bmatrix}$$

For $\theta = 45°$:

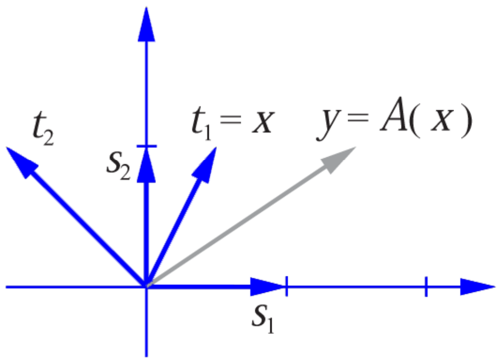
$$\mathbf{A}' = \begin{bmatrix} 5/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

6

Example - (3)

Try a test vector: $\mathbf{x} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$ $\mathbf{x}' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} \qquad \mathbf{y}' = \mathbf{A}'\mathbf{x}' = \begin{bmatrix} 5/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}$$



Check using reciprocal basis vectors:

$$\mathbf{y}' = \mathbf{B}^{-1}\mathbf{y} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}$$

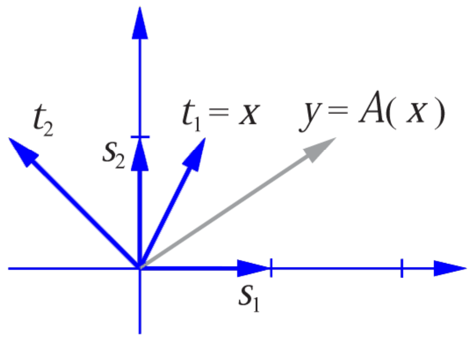
تغییر پایه‌ها

مثال (۳ از ۳)

New Basis Sets

Try a test vector: $\mathbf{x} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$ $\quad \mathbf{x}' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} \qquad \mathbf{y}' = \mathbf{A}'\mathbf{x}' = \begin{bmatrix} 5/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}$$



Check using reciprocal basis vectors:

$$\mathbf{y}' = \mathbf{B}^{-1}\mathbf{y} = \begin{bmatrix} 0.5 & -1 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \end{bmatrix}\begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -2/3 \end{bmatrix}$$

تبدیل‌های خطی برای شبکه‌های عصبی

۴

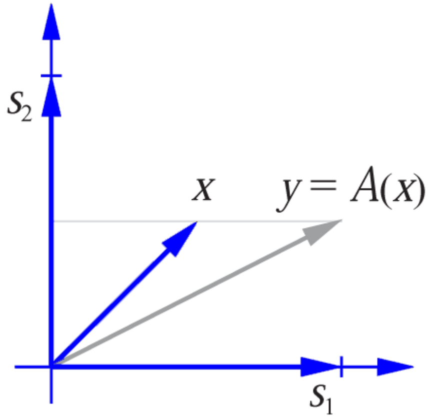
مقدارهای ویژه و بردارهای ویژه

6 Eigenvalues and Eigenvectors

Let $A{:}X \rightarrow X$ be a linear transformation. Those vectors $z \in X$, which are not equal to zero, and those scalars λ which satisfy

$$A(z) = \lambda z$$

are called eigenvectors and eigenvalues, respectively.



Can you find an eigenvector for this transformation?

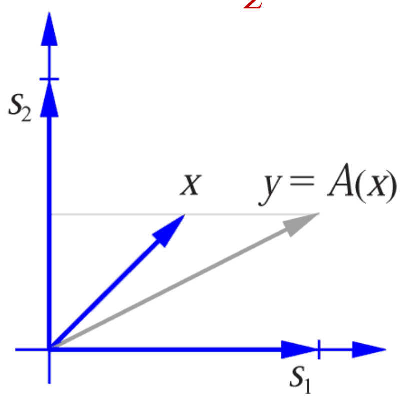
EIGENVALUES AND EIGENVECTORS

Let $A{:}X \rightarrow X$ be a linear transformation. Those vectors $z \in X$, which are not equal to zero, and those scalars λ which satisfy

$$A(z) = \lambda z$$

are called eigenvectors and eigenvalues, respectively.

z λ



Can you find an eigenvector for this transformation?

6 Computing the Eigenvalues

$$\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$$

$$[\mathbf{A} - \lambda\mathbf{I}]\mathbf{z} = \mathbf{0} \quad \Rightarrow \quad |[\mathbf{A} - \lambda\mathbf{I}]| = 0$$

Skewing example (45°):

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \left|\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\right| = 0 \qquad (1-\lambda)^2 = 0 \qquad \begin{matrix} \lambda_1 = 1 \\ \lambda_2 = 1 \end{matrix}$$

$$\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\mathbf{z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\mathbf{z}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} z_{11} \\ z_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{21} = 0 \qquad \mathbf{z}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For this transformation there is only one eigenvector.

محاسبه‌ی مقادیر ویژه

COMPUTING THE EIGENVALUES

$$\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$$

$$[\mathbf{A} - \lambda\mathbf{I}]\mathbf{z} = \mathbf{0} \quad \Rightarrow \quad |[\mathbf{A} - \lambda\mathbf{I}]| = 0$$

Skewing example (45°):

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \left| \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} \right| = 0 \qquad (1-\lambda)^2 = 0 \qquad \begin{matrix} \lambda_1 = 1 \\ \lambda_2 = 1 \end{matrix}$$

$$\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\mathbf{z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\mathbf{z}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} z_{11} \\ z_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{21} = 0 \qquad \mathbf{z}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For this transformation there is only one eigenvector.

MATLAB

```
>> nnd6eg
```

تبدیل‌های خطی برای شبکه‌های عصبی

۵

قطری‌سازی

Diagonalization

6

Perform a change of basis (similarity transformation) using the eigenvectors as the basis vectors. If the eigenvalues are distinct, the new matrix will be diagonal.

$$\mathbf{B} = \begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \dots & \mathbf{z}_n \end{bmatrix}$$

$\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}$ Eigenvectors

$\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ Eigenvalues

$$[\mathbf{B}^{-1}\mathbf{A}\mathbf{B}] = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

DIAGONALIZATION

Perform a change of basis (similarity transformation) using the eigenvectors as the basis vectors. If the eigenvalues are distinct, the new matrix will be diagonal.

با استفاده از بردارهای ویژه به‌عنوان بردارهای پایه، تغییر پایه را انجام می‌دهیم (تبدیل شباهت).
اگر مقادیر ویژه متمایز باشند ⇐ بردارهای ویژه مستقل خطی هستند ⇐ ماتریس جدید قطری خواهد بود.

$$\mathbf{B} = \begin{bmatrix} \mathbf{z}_1 & \mathbf{z}_2 & \dots & \mathbf{z}_n \end{bmatrix}$$

$\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}$ Eigenvectors

$\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ Eigenvalues

$$[\mathbf{B}^{-1}\mathbf{A}\mathbf{B}] = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

Example

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

$$\left| \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \right| = 0 \qquad \lambda^2 - 2\lambda = (\lambda)(\lambda - 2) = 0 \qquad \begin{matrix} \lambda_1 = 0 \\ \lambda_2 = 2 \end{matrix} \qquad \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \mathbf{z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda_1 = 0 \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{z}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{21} = -z_{11} \qquad \mathbf{z}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_2 = 2 \Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{z}_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{22} = z_{12} \qquad \mathbf{z}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Diagonal Form: $\mathbf{A}' = [\mathbf{B}^{-1}\mathbf{A}\mathbf{B}] = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$

قطری‌سازی

مثال

DIAGONALIZATION

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

$$\left| \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \right| = 0 \qquad \lambda^2 - 2\lambda = (\lambda)(\lambda-2) = 0 \qquad \begin{matrix} \lambda_1 = 0 \\ \lambda_2 = 2 \end{matrix} \qquad \begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} \mathbf{z} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda_1 = 0 \Rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{z}_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{21} = -z_{11} \qquad \mathbf{z}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_2 = 2 \Rightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{z}_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad z_{22} = z_{12} \qquad \mathbf{z}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

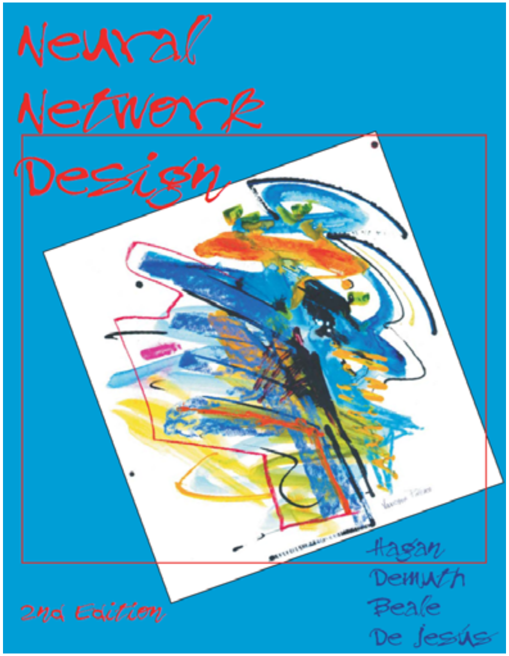
Diagonal Form: $\mathbf{A}' = [\mathbf{B}^{-1}\mathbf{A}\mathbf{B}] = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$

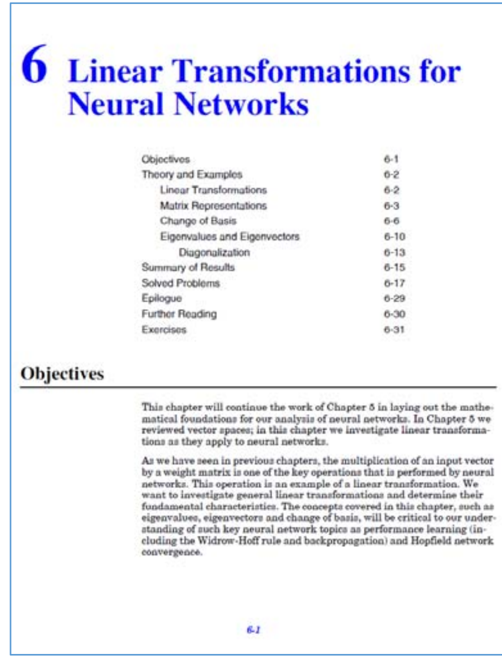
تبدیل‌های خطی برای شبکه‌های عصبی

۶

منابع

منبع اصلی





6 Linear Transformations for Neural Networks

Objectives

This chapter will continue the work of Chapter 5 in laying out the mathematical foundations for our analysis of neural networks. In Chapter 5 we reviewed vector spaces; in this chapter we investigate linear transformations as they apply to neural networks.

As we have seen in previous chapters, the multiplication of an input vector by a weight matrix is one of the key operations that is performed by neural networks. This operation is an example of a linear transformation. We want to investigate general linear transformations and determine their fundamental characteristics. The concepts covered in this chapter, such as eigenvalues, eigenvectors and change of basis, will be critical to our understanding of such key neural network topics as performance learning (including the Widrow-Hoff rule and backpropagation) and Hopfield network convergence.

Martin T. Hagan, Howard B. Demuth, Mark H. Beale, Orlando De Jesus,
Neural Network Design,
2nd Edition, Martin Hagan, 2014.
Chapter 6

Online version can be downloaded from: http://hagan.okstate.edu/nnd.html