Linear Algebra for Machine Learning

1) Basics

- Scaler Quantities
  - Mass
  - Duration
  - length

  (They just have Magnitude)

- Vector Quantities
  - force
  - Velocity
  - Momentum

  (They have both magnitude and Direction)

$\underline{a}$ ⇒ Altı çizgi vektör olduğunu belli eden notation'lardan sadece bir tanesi. ($\bar{a} = (2,1,3)$ veya $\bar{a} = \langle 2,1,3 \rangle$) ile de gösterilebilir.

3 boyutlu (x, y, z)

ÖR

$\langle 4,5 \rangle \rightarrow$ 2 boyutlu (x, y)

2) Vector Dot and Cross Product

Dot Product (scaler product) → Gives a scaler (ordinary number) answer

Cross Product (Vector product) → Gives a vector as an answer

- **Dot Product**

$$\underline{a} = \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix} \qquad \underline{b} = \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$$

$$\underbrace{\underline{a} \cdot \underline{b}}_{\text{Dot Product}} = (-4)\times(3) + (2\times 1) + 1\times 3 = -7 \quad \} \text{ Dot Product Result}$$

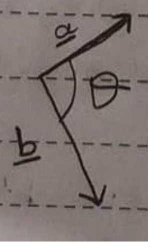
$$\underline{a} \cdot \underline{a} = \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix} \Rightarrow 16 + 4 + 1 = 21$$

$$\hat{\underline{a}} = \frac{1}{\sqrt{21}} \begin{pmatrix} -4 \\ 2 \\ 1 \end{pmatrix} \Rightarrow \hat{\underline{a}} \cdot \hat{\underline{a}} = \frac{1}{21} \times 21 = 1$$

Unit vector ; $\sqrt{21} \Rightarrow$ Magnitude of vector

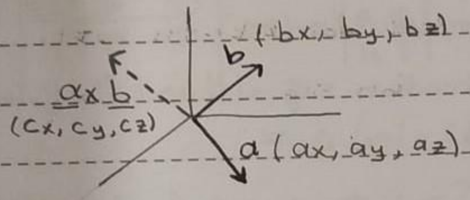
Diğer bir tanım ile Dot Product

↳ $\underline{a} \cdot \underline{b} = |\underline{a}|\,|\underline{b}| \cdot \cos\theta$



• Cross Product

$$\underbrace{\underline{a} \times \underline{b}}_{\text{Cross Product}} = \underline{c}$$



- $c_x = a_y b_z - a_z b_y$
- $c_y = a_z b_x - a_x b_z$
- $c_z = a_x b_y - a_y b_x$

Diğer bir tanım ile Cross Product

↳ $|\underline{a} \times \underline{b}| = |\underline{a}|\,|\underline{b}|\,\sin(\theta)$ → a ile b arasındaki açı

↳ magnitude (length) of vector a

ÖR

$\underline{a} \times \underline{b} = \underline{c}$

$$\begin{pmatrix}2\\3\\4\end{pmatrix} \times \begin{pmatrix}5\\6\\7\end{pmatrix} = \begin{pmatrix}\square\\\square\\\square\end{pmatrix}$$

□ → Bulalım $\begin{pmatrix}\square\\-\\-\end{pmatrix} = \begin{pmatrix}2\\3\\4\end{pmatrix}\begin{pmatrix}5\\6\\7\end{pmatrix} - \begin{pmatrix}2\\3\\4\end{pmatrix}\begin{pmatrix}5\\6\\7\end{pmatrix}$

□ = -3 = 21 - 24

□ → Bulalım $\begin{pmatrix}-\\\square\\-\end{pmatrix} = \begin{pmatrix}2\\3\\4\end{pmatrix}\begin{pmatrix}5\\6\\7\end{pmatrix} - \begin{pmatrix}2\\3\\4\end{pmatrix}\begin{pmatrix}5\\6\\7\end{pmatrix}$

□ = 6 = 20 - 14

□ → Bulalım $\begin{pmatrix}-\\-\\\square\end{pmatrix} = \begin{pmatrix}2\\3\\4\end{pmatrix}\begin{pmatrix}5\\6\\7\end{pmatrix} - \begin{pmatrix}2\\3\\4\end{pmatrix}\begin{pmatrix}5\\6\\7\end{pmatrix}$

□ = -3 = 12 - 15

* $\underline{c} = \begin{pmatrix}-3\\6\\-3\end{pmatrix}$

↳ Hem $\underline{a}$'ya hem $\underline{b}$'ye **Diktir.**

⇓

* $\underline{a} \cdot \underline{c} = 0$

* $\underbrace{\underline{b} \cdot \underline{c}}_{\text{Dot Products.}} = 0$

Özet Cross Product

$\underline{c} = \underline{a} \times \underline{b}$

1) Magnitude of $\underline{c} \Rightarrow |\underline{c}| = |\underline{a}| \cdot |\underline{b}| \cdot \sin\theta$

2) Direction of $\underline{c}$ is perpendicular (Dik) to both $\underline{a}$ and $\underline{b}$. (Direction sağ el kuralı ile bulunur)

3) $\underline{a} \times \underline{b} = -\underline{b} \times \underline{a}$

3) Matrix and Matrix Product

$\underline{\underline{A}} \Rightarrow$ A matrisi

- Matris Çarpımı:

$\underline{\underline{A}}\,\underline{\underline{B}} = \underline{\underline{C}}$ $\quad \underline{\underline{A}} = \begin{bmatrix} 3 & 2 \\ 0 & 3 \\ -1 & 4 \end{bmatrix}$ $\quad \underline{\underline{B}} = \begin{bmatrix} 1 & 0 \\ 2 & -3 \end{bmatrix}$

★ Multiply each row of A by each column of B. ★

$\underline{\underline{C}} = \begin{bmatrix} 7 & -6 \\ 6 & -9 \\ 7 & -12 \end{bmatrix}$

- Determinant :

- Determinant is a Scaler (single) number derived from a Square Matrix.

Determinant of M is $|\underline{\underline{M}}|$

$\underline{\underline{M}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\quad |\underline{\underline{M}}| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$

Determinant $\Rightarrow \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

• 3x3 determinant is evaluated by breaking it into Several 2x2 Determinants.

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

Bu örnekte bu satırı seçmiş olalım. (Herhangi bir sütun, satır olabilir, aynı sonucu verir)

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a(ei-fh) - b(di-gf) + c(dh-ge)$$

What is Determinant for?

Help us to find the inverse of a Matrix, tells us things about the Matrix that are useful in Systems of linear equations, calculus and more.

4) Eigenvalues and Eigenvectors

$A \cdot v = \lambda \cdot v$ → sabit sayı = {3, 5, ...}

A: $n \times n$ kare matris; v: $n \times 1$ sütun matris

$A \cdot v = \lambda \cdot v$ (λ: eigenvalue, v: eigenvector)

Bu λ ve V A matrisinin özdeğer ve özvektörüdür.

$A \cdot v = \lambda \cdot v \Rightarrow A \cdot v - \lambda \cdot v = 0 \Rightarrow (A - \lambda) \cdot v = 0$

sbt bir sayı. Matristen sabit sayı çıkmaz. Bu yüzden sabit sayıyı matris yapmamız lazım.

$(A - \lambda I) \cdot v = 0$

Birim matris $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

★ $|A - \lambda I| = 0 \Rightarrow$ Bu determinantı 0 yapan değerler özdeğerlerdir.

★ $(A - \lambda_n I) \cdot v = 0 \Rightarrow$ Bu eşitliği sağlayan V'ler ise özvektörlerdir.

ÖR →

$A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$ matrisinin özdeğerlerini ve özvektörlerini bulalım.

1. Adım $|A - \lambda I| = 0$

$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ $\lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$ $A - \lambda I = \begin{bmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{bmatrix}$

$\begin{vmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - \lambda - 6 = 0$ $\underbrace{\lambda = \{3, -2\}}_{\text{özdeğerler}}$

2. Adım (Her λ için V (özvektör) bul)

$\lambda = 3$ için;

V (sütun matris)

$(A - 3I) \cdot V = 0 \Rightarrow \begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$-x_1 - 4x_2 = 0 \longrightarrow x_2 = a$; $x_1 = -4a$

$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a \underbrace{\begin{bmatrix} -4 \\ 1 \end{bmatrix}}_{\text{özvektör when } \lambda = 3} \Rightarrow V = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$

★ Aynı işlemi $\lambda = -2$ için de yaparız.

Özellikler

1) $\det A = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n$

2) $A \cdot V = \lambda \cdot V$

3) Özdeğerlerin toplamı Matrisin köşegendeki elemanlarının toplamına eşittir.

Dipnot

A scalar is a number like 3, -5, 0 etc

A vector is a list of numbers (can be in a row or column)

A matrix is an array of numbers (one or more rows, one or more columns)

In fact; vector is also matrix!

Scaler	Vector (row or column)	Matrix (row(s) x column(s))
24	$\begin{bmatrix} 2 & -8 & 7 \end{bmatrix}$ $\begin{bmatrix} 2 \\ -8 \\ 7 \end{bmatrix}$	$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}$

★ If a matrix has only one Row or only one column it is called a Vector

Not:

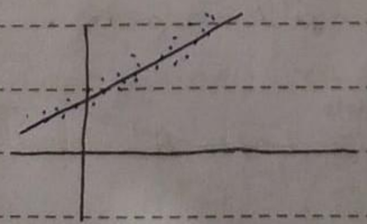
Eigenvectors and eigenvalues have many important applications in computer vision and ML. Well known examples are PCA and Eigen Faces for face recognition

5) Least Square Approximation

Data

X	x_1	x_2	...	x_d
Y	y_1	y_2	...	y_d

Problem: find a "best fit" straight line of the form $y = mx + b$



$$\left.\begin{aligned} y_1 &= mx_1 + b \\ y_2 &= mx_2 + b \\ &\vdots \\ y_d &= mx_d + b \end{aligned}\right\} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_d & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$$

$$\star\ \vec{y} = A\vec{S}$$

../..../....

$$\vec{w} = m(A\vec{e}_1) + b(A\vec{e}_2) \Big\} \quad A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_d & 1 \end{bmatrix}$$

A'nın ilk column'u

We are trying to minimize;

$$\hookrightarrow \|\vec{y} - (m(A\vec{e}_1) + b(A\vec{e}_2))\|$$

Square

$$= \underbrace{\sum_{i=1}^{d} (y_i - (mx_i + b))^2}_{\text{we want to minimize this expression}}$$

To solve that equation, we apply theorem given below;

$$\Rightarrow A^T A \vec{s} = A^T \vec{y}$$

$$\vec{s} = \begin{bmatrix} m \\ b \end{bmatrix}$$

$$A^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \\ 1 & 1 & \cdots & 1 \end{bmatrix} \qquad A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_d & 1 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix}$$

* Then Solve $A^T A \vec{s} = \vec{A} \vec{y}$

$$A^T A = \begin{bmatrix} \sum_{i=1}^{d} x_i^2 & \sum_{i=1}^{d} x_i \\ \sum_{i=1}^{d} x_i & d \end{bmatrix} \qquad A^T \vec{y} = \begin{bmatrix} \sum_{i=1}^{d} x_i y_i \\ \sum_{i=1}^{d} y_i \end{bmatrix}$$

Not:

(Matrisin Tersini Bulma) (Elementer Satır İşlemleri Metodu)

$$[A \,\vdots\, I] \longrightarrow [I \,\vdots\, A^{-1}]$$

A: Tersi alınacak matris

I: Birim matris

Birim matris yapmak (Amaç)

A'yı birim matris yaptığımızda diğer tarafta A^{-1} oluşur.

1.adım: Genişletilmiş Katsayılar Matrisi oluştur.

2.olarak 1 yapılmalılar

3.olarak 0 yapılmalı

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ olsun } \Rightarrow \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \longrightarrow \text{Genişletilmiş Katsayılar Matrisi}$$

1 olarak 0 yapılmalı

Burayı I yapmalıyız ki A^{-1} oluşsun.

2. Adım:

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \xrightarrow{-3S_1 + S_2 \to S_2} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right] \xrightarrow{-1/2 S_2 \to S_2}$$

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right] \xrightarrow{-2S_2 + S_1 \to S_1} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right]$$

$$A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$$

Devam :)

$$\begin{bmatrix} m \\ b \end{bmatrix} = \vec{s} = (A^T A)^{-1} (A^T \vec{y}) \Rightarrow \text{if } (A^T A)^{-1} \text{ exists.}$$

2x2 matriste Tersini alma kolay yol;

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \underbrace{\frac{1}{ad-bc}}_{\text{(Determinant)}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Unutma!! $A \times A^{-1} = I$

Neden Ters alma işlemine gerek duyduk?

Matrislerde bölüm yapamayız. Yani $XA = B$ iken 2 tarafı da A'ya bölüp $X = B/A$ yapamayız.

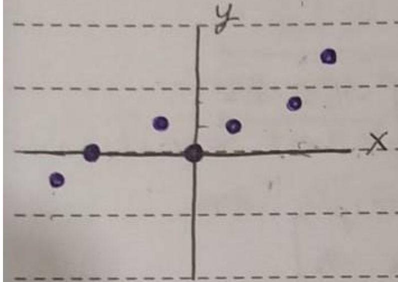
Bunun yerine; $XAA^{-1} = BA^{-1} \Rightarrow XI = BA^{-1} \Rightarrow X = BA^{-1}$ olur.

$$(A^T A)^{-1} (A^T \vec{y}) = \begin{bmatrix} m \\ b \end{bmatrix}$$

Sonuç olarak;

$$\begin{bmatrix} m \\ b \end{bmatrix} = \frac{1}{d\sum_i x_i^2 - \sum_{i,j} x_i x_j} \begin{bmatrix} d\sum_l x_l y_l & -\sum_{k,l} x_k y_l \\ \sum_{k,l} x_k^2 y_l & -\sum_{k,l} x_k x_l y_l \end{bmatrix}$$

ÖR:



$y = mx + b$

$$A = \begin{array}{c} \begin{matrix} x & 1 \end{matrix} \\ \begin{bmatrix} -4 & 1 \\ -3 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \end{array}$$

$$\vec{y} = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 2 \\ 4 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 52 & 0 \\ 0 & 7 \end{bmatrix}$$

$$\left.\begin{array}{l} A^T \vec{y} = \begin{bmatrix} 26 \\ 7 \end{bmatrix} \Rightarrow \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} \end{array}\right\} y = 1/2\, x + 1$$

6) Linear Combinations of Vectors

(scaler)(something 1) ± (scaler)(something 2) ± (scaler)(something 3)

ÖR $\vec{v_1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}$, $\vec{v_3} = \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$

The vector $\vec{b} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$ is a linear combination of $\vec{v_1}, \vec{v_2}, \vec{v_3}$

Why is this True?

This vector can be written as a combination of three given vectors using scaler multiplication and addition

$$\begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 0\begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$$

$$* \boxed{\vec{b} = 3\vec{v_1} + 0\vec{v_2} + 0\vec{v_3}}$$

!! Genel olarak, bu 3 vektörün Tüm doğrusal Kombinasyonları, bunların "Span"ı olarak refer edilir
↳ Written as Span($\vec{v_1}, \vec{v_2}, \vec{v_3}$). önceki sayfadaki $\begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$ are member of this set.

Formal Definition: Let the vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}, \dots, \vec{v_n}$ be vectors in R^n and $c_1, c_2, \dots, c_n$ be scalers. Then the vector $\vec{b} = c_1\vec{v_1} + c_2\vec{v_2} + \dots + c_n\vec{v_n}$ is called a linear combination of $\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$. The scalers $c_1, c_2, \dots, c_n$ are called "weights".

7) Linear Transformations (Doğrusal Dönüşümler)

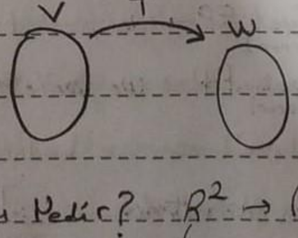
Doğrusal dönüşümler fonksiyonlardır.

$T(v) = \dots$

Girdi — çıktı

Dönüşüm fonksiyonları: $T(v) = w$

↳ Bir vektör uzayından diğerine tanımlanırlar.



Vektör uzayı nedir? $R^2 \rightarrow \binom{1}{0}, \binom{2}{1}$ veya $R^3 \rightarrow \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$

(2 boyutlu vektör uzayı)

Dönüşüm fonk $\Rightarrow$ Girdileri R^2'den R^3'e dönüştürür (örneğin)

ÖR $T: R^3 \to R \longrightarrow T(x_1, x_2, x_3) = (x_1 \cdot x_2 \cdot x_3)$

$T: R^2 \to R^3 \longrightarrow T(x_1, x_2) = (x_1 + x_2, x_1 - x_2, x_1 \cdot x_2)$

Bir dönüşüm fonk. Doğrusal dönüşüm fonk olabilmesi için;

1) $T(u+v) = T(u) + T(v)$

2) $c \in R$ için; $T(c \cdot u) = c\, T(u)$

Sağlamalıdır.

$V \xrightarrow{T} W$

$u \in V$

$v \in V$