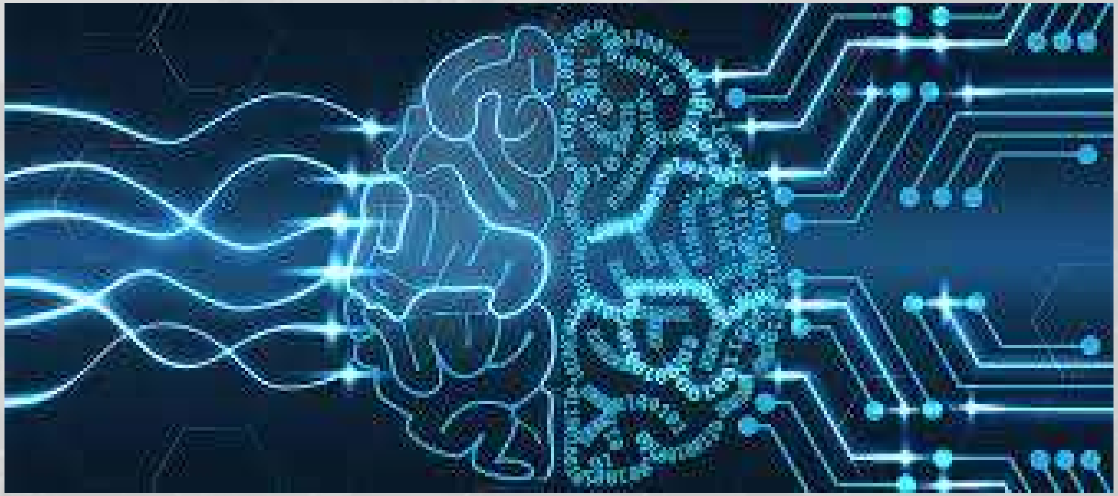
14620323
DEEP LEARNING



Mathematics for Machine Learning

PENGAMPU

Dr. Fajar Astuti Hermawati, S.Kom.,M.Kom.

Elsen Ronando, S.Si.,M.Si

Bagus Hardiansyah, S.Kom.,M.Si

Andrey Kartika Widhy H., S.Kom., M.Kom.

Capaian Pembelajaran

- Mampu mengidentifikasi konsep matematika dan mesin pembelajaran dasar untuk algoritma deep learning. [C2, A3]

Bahan Kajian

- Linear Algebra
- Probability and InformationTheory
- Numerical Computation

Linear Algebra

Scalars

- Skalar adalah bilangan tunggal
- Bilangan bulat, bilangan real, bilangan rasional, dll.
- Kita menunjukkannya (notasi) dengan huruf miring : a, n, x
  - “Misalkan $s \in \mathbb{R}$ adalah kemiringan garis,” ketika mendefinisikan skalar bernilai riil
  - “Misalkan $n \in \mathbb{N}$ menjadi jumlah unit” saat mendefinisikan skalar bilangan asli

Vectors

- Vektor adalah larik (array) angka 1-D

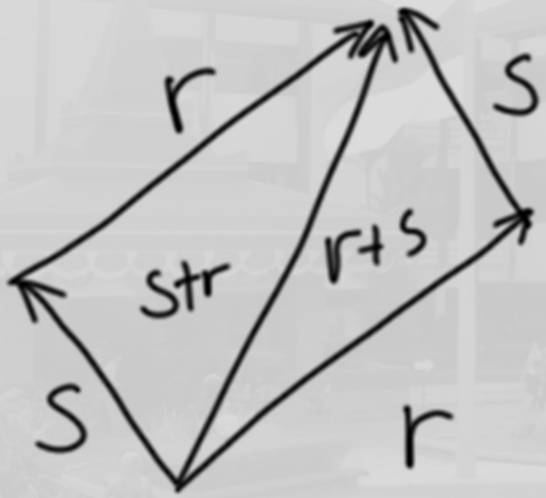
$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

- Bisa real, biner, integer, dll.
- Jika setiap elemen dalam $\mathbb{R}$, dan vektor memiliki n elemen, maka vektor terletak pada himpunan yang dibentuk dengan mengambil produk Cartesian dari $\mathbb{R}$ n kali, dilambangkan dengan $\mathbb{R}^n$

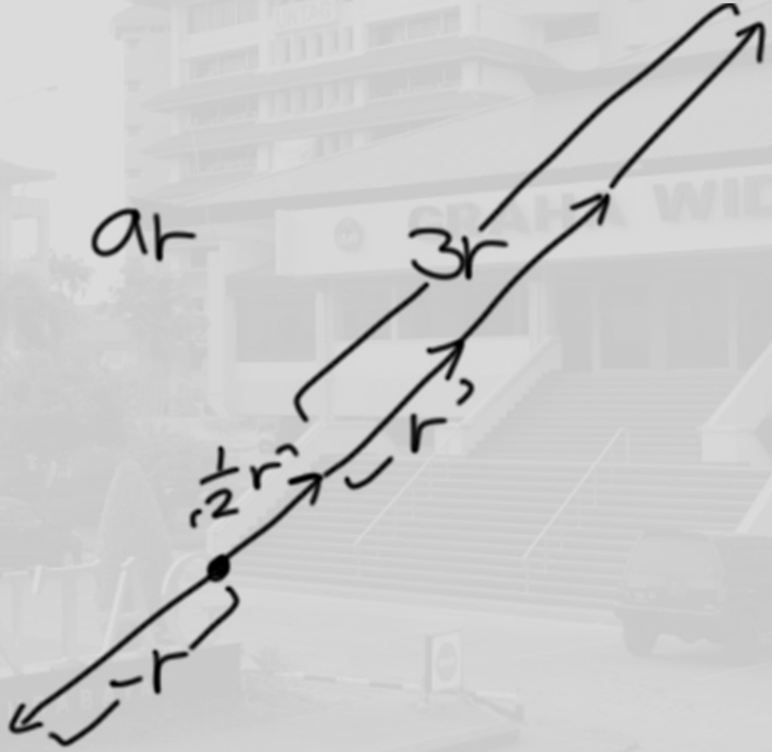
Vectors

- Terkadang kita perlu mengindeks sekumpulan elemen vektor.
- Dalam hal ini, kita mendefinisikan himpunan yang berisi indeks dan menulis himpunan sebagai subskrip.
  - Misalnya, untuk mengakses x_1, x_3 dan x_6, kita mendefinisikan himpunan $S = \{1, 3, 6\}$ dan menulis x_S.
- Kita menggunakan tanda − untuk mengindeks komplemen dari suatu himpunan.
  - Misalnya x_{-1} adalah vektor yang memuat semua elemen x kecuali x_1, dan x_{-S} adalah vektor yang memuat semua elemen x kecuali x_1, x_3 dan x_6.

Vector



$r+s = s+r$



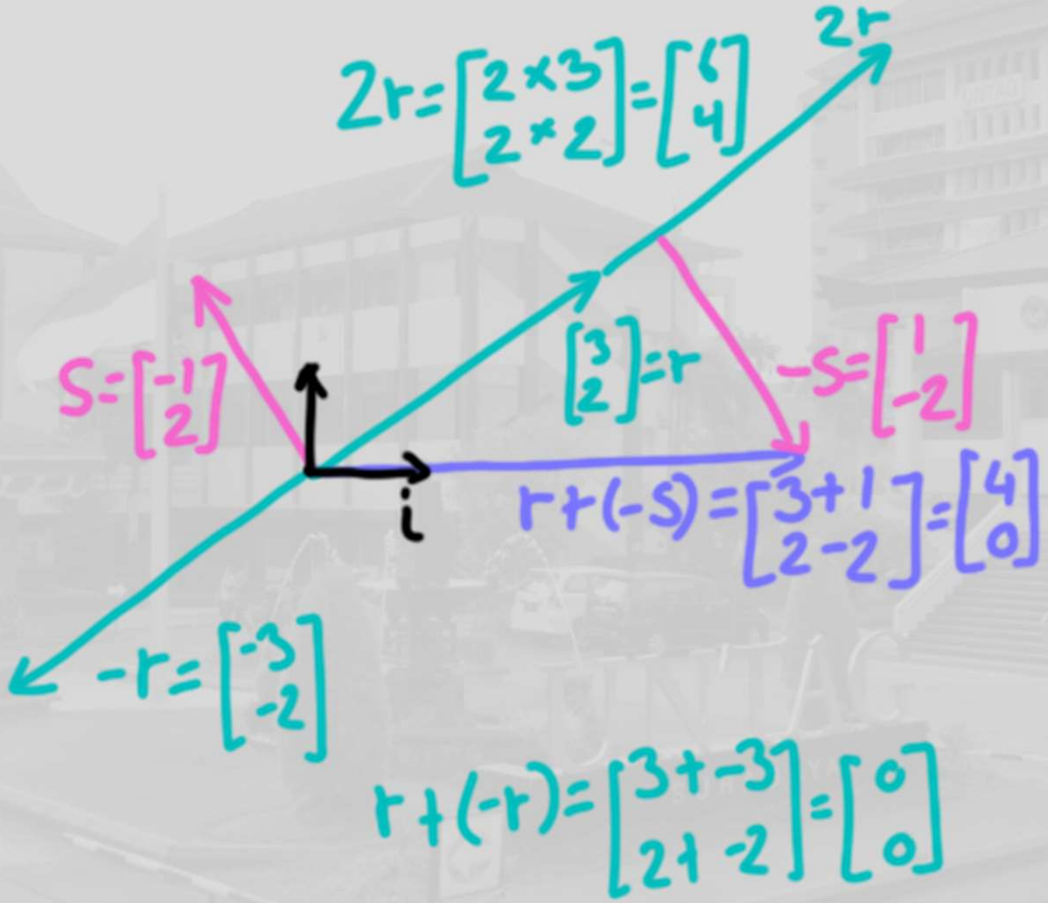
Contoh Vector

$$\begin{array}{l} \begin{bmatrix} 120 \\ 2 \\ 1 \\ 150 \end{bmatrix} \end{array} \begin{array}{l} m^2 \\ \text{kamar tidur} \\ \text{toilet} \\ \text{juta} \end{array}$$

$+$

$$2\begin{bmatrix} 120 \\ 2 \\ 1 \\ 150 \end{bmatrix} = \begin{bmatrix} 240 \\ 4 \\ 2 \\ 300 \end{bmatrix} \begin{array}{l} m^2 \\ \text{kamar tidur} \\ \text{toilet} \\ \text{juta} \end{array}$$

Vector Operations



Rules of Vector

commutative

↳ $s.r = r.s = r_i.s_i + r_j.s_j$

$= 3 \times -1 + 2 \times 2 = -3+4 = 1$

distributive

↳ $r(s+t) = r.s + r.t$

associative

↳ $r.(as) = a.(rs)$

$r_i(a s_i) + r_j(a.s_j) = a(r_i s_i + r_j s_j)$

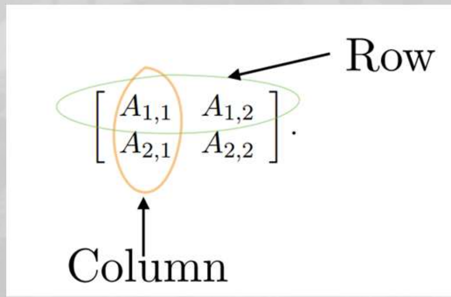
Length

$$r = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} r_i \\ r_j \end{bmatrix}$$

$$r \cdot r = r_i r_i + r_j r_j$$
$$= r_i^2 + r_j^2$$
$$r \cdot r = |r|^2$$

Matrices

- Matriks adalah susunan (array) angka 2-D



- Kita biasanya memberi nama variabel matriks huruf besar dengan huruf tebal, seperti ***A***.
- Jika matriks bernilai nyata ***A*** memiliki tinggi m dan lebar n, maka kita mengatakan bahwa $\boldsymbol{A} \in \mathbb{R}^{m \times n}$

Matrices

- Kita biasanya mengidentifikasi elemen matriks menggunakan namanya dalam huruf miring tetapi bukan huruf tebal, dan indeksnya dicantumkan dengan koma pemisah.
  - Misalnya, $A_{1,1}$ adalah entri kiri atas $\boldsymbol{A}$ dan Am,n adalah entri kanan bawah.
- Kita dapat mengidentifikasi semua angka dengan koordinat vertikal *i* dengan menulis ":" untuk koordinat horisontal.
  - Misalnya, $A_{i,:}$ menunjukkan penampang horizontal $\boldsymbol{A}$ dengan koordinat vertikal *i*. Ini dikenal sebagai baris ke-*i* dari $\boldsymbol{A}$. Demikian juga, $\boldsymbol{A}_{:,i}$ adalah kolom ke-*i* dari $\boldsymbol{A}$.

Matrices

- Ketika kita perlu mengidentifikasi secara eksplisit elemen-elemen suatu matriks, kita menuliskannya sebagai larik yang diapit tanda kurung siku:

$$\begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}.$$

- Terkadang kita mungkin perlu mengindeks ekspresi bernilai matriks yang bukan hanya satu huruf. Dalam hal ini, kita menggunakan subskrip setelah ekspresi tetapi tidak mengubah apa pun menjadi huruf kecil.
  - Misalnya, $f(\boldsymbol{A})_{i,j}$ memberikan elemen (i,j) dari matriks yang dihitung dengan menerapkan fungsi f ke $\boldsymbol{A}$

Tensors

- Tensor adalah larik (array) angka, yang mungkin ada
  - nol dimensi, dan menjadi skalar
  - satu dimensi, dan menjadi vektor
  - dua dimensi, dan menjadi matriks
  - atau lebih dimensi
- Kita menganotasi tensor bernama "A" dengan jenis huruf ini: **A**.
- Kita mengidentifikasi elemen **A** pada koordinat (*i*, *j*, *k*) dengan menulis $\mathbf{A}_{i,j,k}$.

Matrix Transpose

- Transpose matriks adalah bayangan cermin dari matriks melintasi garis diagonal, yang disebut diagonal utama, berjalan ke bawah dan ke kanan, dimulai dari sudut kiri atas.

$$\boldsymbol{A} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{bmatrix} \Rightarrow \boldsymbol{A}^\top = \begin{bmatrix} A_{1,1} & A_{2,1} & A_{3,1} \\ A_{1,2} & A_{2,2} & A_{3,2} \end{bmatrix}$$

- Kita menyatakan transpose matriks ***A*** sebagai $\mathbf{A}^\top$, dan didefinisikan sedemikian rupa $(\boldsymbol{A}^\top)_{i,j} = A_{j,i}.$ dan berlaku $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top.$

add matrices

- dengan menjumlahkan elemen-elemennya yang bersesuaian: $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ di mana $C_{i,j} = A_{i,j} + B_{i,j}$
- Kita juga dapat menambahkan skalar ke matriks atau mengalikan matriks dengan skalar, hanya dengan melakukan operasi tersebut pada setiap elemen matriks: $\boldsymbol{D} = a \cdot \boldsymbol{B} + c$ di mana $D_{i,j} = a \cdot B_{i,j} + c$

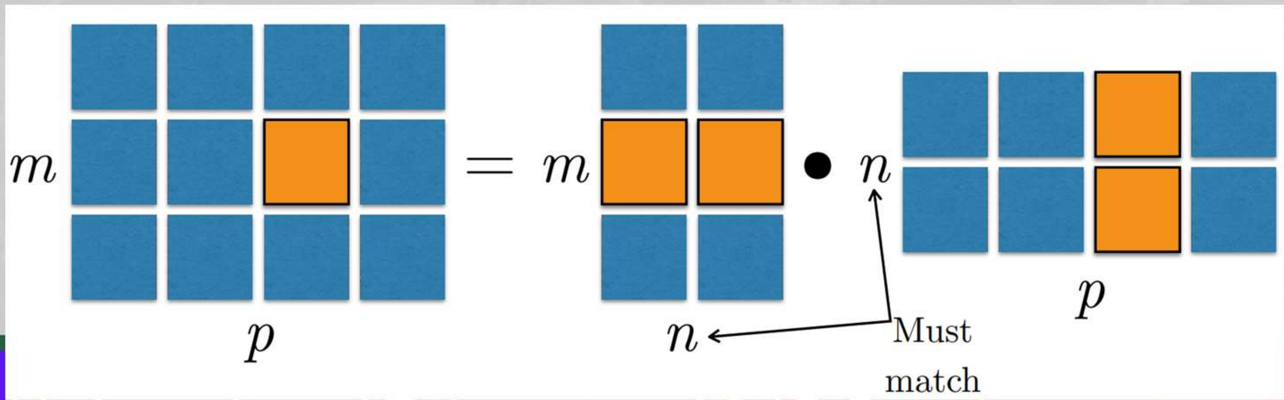
add matrices

- Dalam konteks deep learning, kita juga menggunakan beberapa notasi yang kurang konvensional.
- Kita mengizinkan penambahan matriks dan vektor, menghasilkan matriks lain: $\boldsymbol{C} = \boldsymbol{A} + b$, di mana $C_{i,j} = A_{i,j} + b_j$.
- Dengan kata lain, vektor *b* ditambahkan ke setiap baris matriks.
- Penyingkatan ini menghilangkan kebutuhan untuk mendefinisikan matriks dengan *b* disalin ke setiap baris sebelum melakukan penjumlahan.
- Penyalinan implisit *b* ke banyak lokasi ini disebut ***broadcasting***.

Matrix (Dot) Product

- Perkalian dua buah matriks ***A*** dan ***B*** dan hasilnya disimpan menjadi matriks baru ***C*** dapat dituliskan: ***C*** = ***AB***
- Operasi produk ditentukan oleh

$$C_{i,j} = \sum_k A_{i,k} B_{k,j}.$$



Matrix (Dot) Product

- Perhatikan bahwa produk standar dari dua matriks bukan hanya sebuah matriks yang mengandung produk dari masing-masing elemen.
- Operasi semacam itu ada dan disebut ***element-wise product***, atau produk **Hadamard**, dan dilambangkan sebagai $\boldsymbol{A} \odot \boldsymbol{B}$
- Hasil kali titik antara dua vektor x dan y dengan dimensi yang sama adalah hasil kali matriks $x^{T}y$.

Identity Matrix

- Matriks identitas adalah matriks yang tidak mengubah vektor apa pun ketika kita mengalikan vektor itu dengan matriks itu.
- Kita menunjukkan matriks identitas yang mempertahankan vektor n-dimensi sebagai $\boldsymbol{I}_n$. Secara formal, $\boldsymbol{I}_n \in \mathbb{R}^{n \times n}$, dan

$$\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{I}_n \boldsymbol{x} = \boldsymbol{x}.$$

- Example identity matrix: $\boldsymbol{I}_3$.

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

Systems of Equations

- Kita sekarang cukup mengetahui notasi aljabar linier untuk menuliskan sistem persamaan linier:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

- di mana $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ adalah matriks yang diketahui, $\boldsymbol{b} \in \mathbb{R}^{m}$ adalah vektor yang diketahui, dan $\boldsymbol{x} \in \mathbb{R}^{n}$ adalah vektor dari variabel yang tidak diketahui yang ingin kita selesaikan.
- Setiap elemen x_i dari $\boldsymbol{x}$ adalah salah satu dari variabel yang tidak diketahui ini.
- Setiap baris $\boldsymbol{A}$ dan setiap elemen $\boldsymbol{b}$ memberikan kendala lain.

Systems of Equations

- Kita dapat menulis ulang persamaan sebagai:

$$A_{1,:}x = b_1$$
$$A_{2,:}x = b_2$$
$$\dots$$
$$A_{m,:}x = b_m$$

- atau bahkan lebih eksplisit sebagai

$$A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n = b_1$$
$$A_{2,1}x_1 + A_{2,2}x_2 + \cdots + A_{2,n}x_n = b_2$$
$$\dots$$
$$A_{m,1}x_1 + A_{m,2}x_2 + \cdots + A_{m,n}x_n = b_m.$$

Solusi

$$
\begin{aligned}
x + y + z &= 6 \\
2y + 5z &= -4 \\
2x + 5y - z &= 27
\end{aligned}
$$

$$
\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix}
$$

$AX = B$

- A adalah koefisien x, y, dan z
- X adalah **x, y and z**
- B adalah **6, −4 and 27**

$X = A^{-1}B$

$$
\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 5 \\ 2 & 5 & -1 \end{bmatrix}^{-1} = \frac{1}{-21} \begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix}
$$

Solusi

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{-21}\begin{bmatrix} -27 & 6 & 3 \\ 10 & -3 & -5 \\ -4 & -3 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -4 \\ 27 \end{bmatrix} = \frac{1}{-21}\begin{bmatrix} -105 \\ -63 \\ 42 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$$

$x = 5, y = 3, z = -2$

Solving Systems of Equations

- Sistem persamaan linear dapat memiliki:
  - Tidak ada solusi
  - Banyak solusi
  - Solusi tepat satu: ini berarti perkalian dengan matriks adalah fungsi yang dapat dibalik (invers matrix)

Matrix Inversion

- Matrix inverse dinotasikan $\boldsymbol{A}^{-1}$:

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_n.$$

- Memecahkan sistem menggunakan invers:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$
$$\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$
$$\boldsymbol{I}_n\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$$

- Note: Secara numerik tidak stabil, tetapi berguna untuk analisis abstrak

Inverse Matrix

$$A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$$

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Inverse = (1/determinant)*Adjoint

Inverse Matrix

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{11} \\ a_{32} & a_{31} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}.$$

Determinant A = |A| = a(ei − fh) − b(di − fg) + c(dh − eg)

Invertibilitas

- Matriks tidak dapat dibalik jika…
  - Lebih banyak baris daripada kolom
  - Lebih banyak kolom daripada baris
  - Baris/kolom redundan (“tergantung linier”, “peringkat rendah”)

Norms

- Terkadang kita perlu mengukur ukuran vektor.
- Dalam pembelajaran mesin, kita mengukur ukuran vektor menggunakan fungsi yang disebut ***norm***.
- Secara formal, norm L^p diberikan oleh

$$||\boldsymbol{x}||_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}}$$

- Untuk $p \in \mathbb{R}$, $p \geq 1$

Norms

- ***Norm***, termasuk ***norm*** *L^p*, adalah fungsi yang memetakan vektor ke nilai non-negatif.
- Pada tingkat intuitif, ***norm*** vektor ***x*** mengukur jarak dari titik asal ke titik ***x***.
- Lebih tepatnya, ***norm*** adalah setiap fungsi *f* yang memenuhi sifat-sifat berikut:
  - $f(\boldsymbol{x}) = 0 \Rightarrow \boldsymbol{x} = \boldsymbol{0}$
  - $f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$ (the **triangle inequality**)
  - $\forall \alpha \in \mathbb{R}, f(\alpha \boldsymbol{x}) = |\alpha| f(\boldsymbol{x})$

Norms

- ***Norm*** L^2, dengan $p = 2$, dikenal sebagai ***norm Euclidean***, yang merupakan jarak Euclidean dari titik asal ke titik yang diidentifikasi oleh $\boldsymbol{x}$.
- ***Norm*** L^2 sangat sering digunakan dalam pembelajaran mesin sehingga sering dilambangkan hanya sebagai $||\boldsymbol{x}||$, dengan subskrip 2 dihilangkan.
- Juga umum untuk mengukur ukuran vektor menggunakan norma L^2 kuadrat, yang dapat dihitung secara sederhana sebagai $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

Norms

- Norm L^1 dapat disederhanakan menjadi:

$$||\boldsymbol{x}||_1 = \sum_i |x_i|.$$

- Norm L^1 umumnya digunakan dalam pembelajaran mesin ketika perbedaan antara elemen nol dan bukan nol sangat penting.
- Setiap kali elemen $\boldsymbol{x}$ menjauh dari 0 sebesar ϵ, norm L_1 bertambah sebesar ϵ.

Norms

- Kita terkadang mengukur ukuran vektor dengan menghitung jumlah elemen bukan nolnya.
- Beberapa penulis menyebut fungsi ini sebagai "norm L^0", tetapi ini adalah terminologi yang salah.
- Jumlah entri bukan nol dalam vektor bukanlah ***norm***, karena penskalaan vektor dengan α tidak mengubah jumlah entri bukan nol.
- Norm L^1 sering digunakan sebagai pengganti jumlah entri bukan nol

Norms

- Satu norma lain yang biasa muncul dalam pembelajaran mesin adalah norma L^∞, juga dikenal sebagai ***max norm***.
- Norma ini disederhanakan menjadi nilai absolut dari elemen dengan besaran terbesar dalam vektor,

$$||\boldsymbol{x}||_\infty = \max_i |x_i|.$$

Norms

- Terkadang kita juga ingin mengukur ukuran matriks.
- Dalam konteks pembelajaran mendalam, cara paling umum untuk melakukannya adalah dengan ***norm Frobenius*** yang tidak jelas:

$$||A||_F = \sqrt{\sum_{i,j} A_{i,j}^2},$$

- yang analog dengan norma L^2 vektor

Special Matrices and Vectors

- Vektor satuan (unit vector) adalah vektor dengan norma satuan:

$$||\boldsymbol{x}||_2 = 1.$$

- Matriks simetris adalah matriks apa pun yang sama dengan transposnya sendiri:

$$\boldsymbol{A} = \boldsymbol{A}^\top.$$

- Matriks ortogonal adalah matriks bujur sangkar yang baris-barisnya saling ortonormal dan kolom-kolomnya saling ortonormal:

$$\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I}.$$

- Yang menyebabkan $\boldsymbol{A}^{-1} = \boldsymbol{A}^\top,$

Special Matrices and Vectors

- Sebuah vektor ***x*** dan vektor ***y*** saling ortogonal jika $\boldsymbol{x}^\top \boldsymbol{y} = 0$.
- Jika kedua vektor memiliki norma bukan nol, ini berarti keduanya membentuk sudut 90 derajat satu sama lain.
- Dalam $\mathbb{R}^n$, paling banyak *n* vektor dapat saling ortogonal dengan norma bukan nol.
- Jika vektor tidak hanya ortogonal tetapi juga memiliki norma satuan, kita menyebutnya ***ortonormal***.

Orthogonal Matrices

column vectors $v_1, v_2, \ldots, v_n$ *of* Q *are orthogonal :*

$$v_i^T v_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

The orthogonal matrix $Q = [v_1, v_2, \ldots, v_n]$

$$Q^T \cdot Q = \begin{bmatrix} v_1^T \\ v_2^T \\ \ldots \\ v_n^T \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} = \begin{bmatrix} v_1^T v_1 & v_1^T v_2 & \ldots & v_1^T v_n \\ v_2^T v_1 & v_2^T v_2 & \ldots & v_2^T v_n \\ \ldots & & & \\ v_n^T v_1 & v_n^T v_2 & \ldots & v_n^T v_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & & & \\ 0 & 0 & \ldots & 1 \end{bmatrix} = I$$

$$\therefore Q^{-1} = Q^T$$

Orthogonal Matrices

A Square matrix 'A' is orthogonal if

$$A^T = A^{-1}$$

(OR)

$$AA^T = A^TA = I$$, where

- A^T = Transpose of A
- A^{-1} = Inverse of A
- I = Identity matrix of same order as 'A'

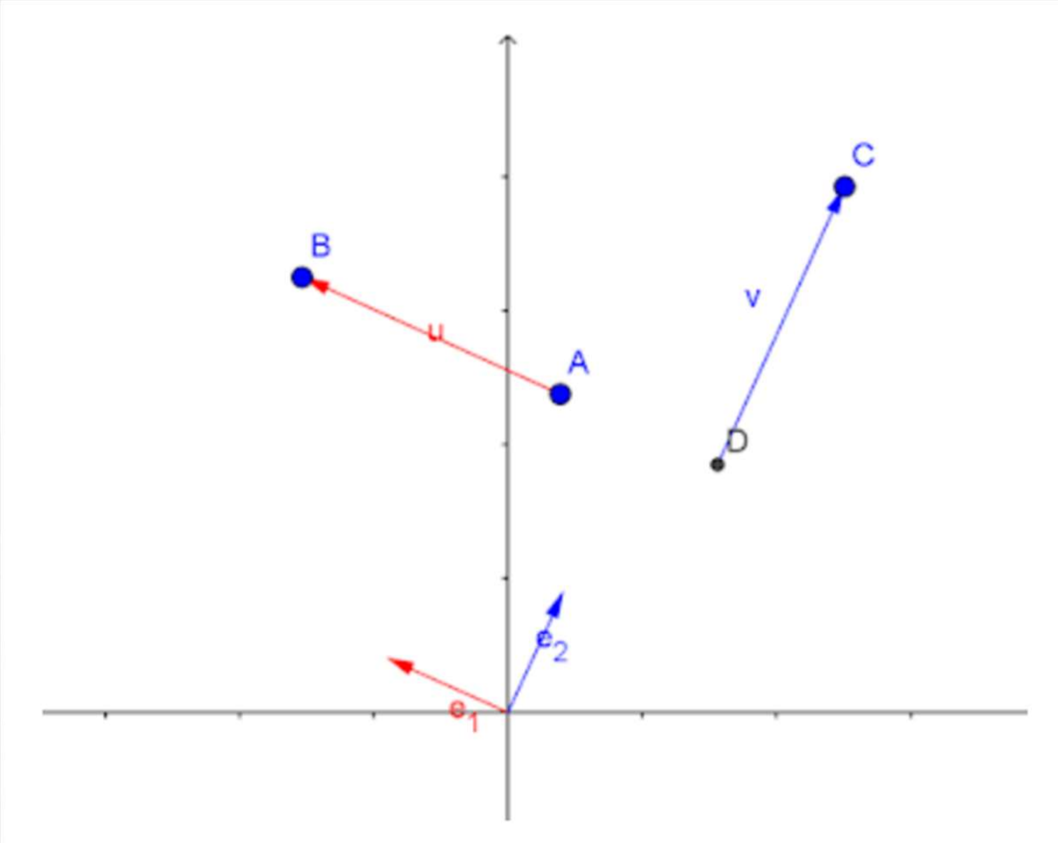
Orthogonal Matrices

- Himpunan vektor-vektor dalam ruang hasilkali dalam disebut sebagai **Himpunan Ortogonal**, jika setiap pasangan vektor yang berbeda dalam himpunan tersebut saling ortogonal.
- Himpunan ortogonal yang setiap vektornya memiliki norma 1 disebut **Ortonormal**.
- Dengan kata lain, {v1 , v2 , … , vn} dari vektor-vektor di **V** adalah ortonormal apabila:
  - <vj , vk> = 0, j != k
  - <vj , vk> = 1, j = k
  - j, k = 1 , 2 , … , n

Orthogonal Matrices

- Suatu himpunan **V** = {**v**1, **v**2, **v**3} dapat dikatakan ortonormal apabila berlaku
  - <**v**1, **v**2> = 0
  - <**v**1, **v**3> = 0
  - <**v**2, **v**3> = 0
  - |**v**1| = |**v**2| = |**v**3| = 1

Orthogonal Matrices



Ortonormal harus memenuhi dua syarat, yaitu panjang vektornya 1 dan saling tegak lurus. Dua vektor yang saling tegak lurus **belum bisa** dikatakan saling ortonormal jika ada vektor yang **panjangnya** tidak sama dengan 1.

Orthogonal Matrices

Example

Diberikan himpunan $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ dengan

$$\mathbf{v}_1 = (0,1,0), \quad \mathbf{v}_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad \mathbf{v}_3 = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

adalah vektor-vektor di R^3 yang dilengkapi hasilkali dalam euclid. Tunjukkan bahwa himpunan $\mathbf{V}$ ortonormal.

Orthogonal Matrices

Solution

- *Himpunan* **V** *dikatakan ortonormal apabila memenuhi*

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0, \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = 0, \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0 \text{ } \textit{dan} \text{ } \|\mathbf{v}_1\| = \|\mathbf{v}_2\| = \|\mathbf{v}_3\| = 1$$

- *Perhatikan bahwa*

$$\begin{aligned}
\langle \mathbf{v}_1, \mathbf{v}_2 \rangle &= 0 \cdot \frac{1}{\sqrt{2}} + 1 \cdot 0 + 0 \cdot \frac{1}{\sqrt{2}} = 0 \\
\langle \mathbf{v}_1, \mathbf{v}_3 \rangle &= 0 \cdot -\frac{1}{\sqrt{2}} + 1 \cdot 0 + 0 \cdot \frac{1}{\sqrt{2}} = 0 \\
\langle \mathbf{v}_2, \mathbf{v}_3 \rangle &= \frac{1}{\sqrt{2}} \cdot -\frac{1}{\sqrt{2}} + 0 \cdot 0 + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = 0
\end{aligned}$$

Orthogonal Matrices

Solution

- *Selanjutnya diperoleh*

$$\begin{aligned}
\|\mathbf{v}_1\| &= \sqrt{0^2 + 1^2 + 0^2} = 1 \\
\|\mathbf{v}_2\| &= \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + 0^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = 1 \\
\|\mathbf{v}_3\| &= \sqrt{\left(-\frac{1}{\sqrt{2}}\right)^2 + 0^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = 1
\end{aligned}$$

- *Dengan demikian* $\mathbf{V}$ *ortonormal.*

Eigendecomposition

- Banyak objek matematika dapat dipahami lebih baik dengan memecahnya menjadi bagian-bagian penyusunnya, atau menemukan beberapa sifat darinya yang bersifat universal, bukan disebabkan oleh cara kita memilih untuk merepresentasikannya.
  - Misalnya, bilangan bulat dapat didekomposisi menjadi faktor prima.
  - Cara kita merepresentasikan angka 12 akan berubah tergantung apakah kita menuliskannya dalam basis sepuluh atau biner, tetapi akan selalu benar bahwa $12 = 2 \times 2 \times 3$.
  - Dari representasi ini kita dapat menyimpulkan sifat-sifat yang berguna, misalnya, bahwa 12 tidak habis dibagi 5, dan kelipatan bilangan bulat apa pun dari 12 akan habis dibagi 3.

Eigendecomposition

- Salah satu jenis dekomposisi matriks yang paling banyak digunakan disebut **dekomposisi eigen**, di mana kita menguraikan matriks menjadi sekumpulan **vektor eigen** dan **nilai eigen**.
- **Vektor eigen** dari matriks bujur sangkar ***A*** adalah vektor bukan nol ***v*** sehingga perkalian dengan ***A*** hanya mengubah skala ***v***:

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

- Skalar λ dikenal sebagai **nilai eigen** yang sesuai dengan vektor eigen ini. (Kita juga dapat menemukan **vektor eigen kiri** sehingga $\boldsymbol{v}^T\boldsymbol{A} = \lambda\boldsymbol{v}^T$, tetapi kita biasanya memperhatikan **vektor eigen kanan**.)

Eigendecomposition

- Misalkan matriks:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}$$

- Mempunyai vector dan nilai eigen:

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \text{with eigenvalue} \quad \lambda_1 = 4$$

$$\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{with eigenvalue} \quad \lambda_2 = -1$$

Eigendecomposition

- Jika $\boldsymbol{v}$ adalah vektor eigen dari $\boldsymbol{A}$, maka vektor apa pun yang diskala ulang $s\boldsymbol{v}$ untuk $s \in \mathbb{R}$, $s \neq 0$.
- Selain itu, $s\boldsymbol{v}$ masih memiliki nilai eigen yang sama. Untuk alasan ini, kita hanya mencari **vektor eigen satuan.**
- Misalkan sebuah matriks $\boldsymbol{A}$ memiliki n vektor eigen bebas linear $\{\boldsymbol{v}^{(1)}, . . . , \boldsymbol{v}^{(n)}\}$ dengan nilai eigen yang sesuai $\{\lambda_1, . . . , \lambda_n\}$.
- Kita dapat menggabungkan semua vektor eigen untuk membentuk matriks $\boldsymbol{V}$ dengan satu vektor eigen per kolom: $\boldsymbol{V} = [v^{(1)}, . . . , v^{(n)}]$.

Eigendecomposition

- Demikian pula, kita dapat menggabungkan nilai eigen untuk membentuk vektor $\lambda = [\lambda_1 , . . .\lambda_n]^T$. Komposisi eigen dari ***A*** dirumuskan:

$$\boldsymbol{A} = \boldsymbol{V} \operatorname{diag}(\boldsymbol{\lambda}) \boldsymbol{V}^{-1}.$$

- Secara khusus, setiap matriks simetris nyata dapat didekomposisi menjadi ekspresi hanya menggunakan vektor eigen dan nilai eigen bernilai nyata:

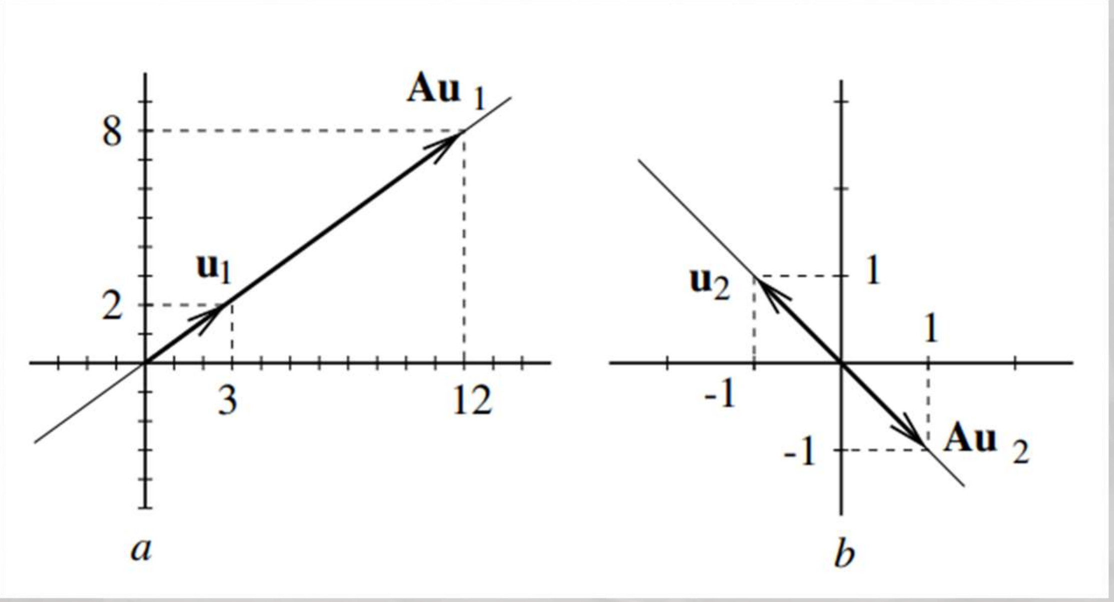
$$\boldsymbol{A} = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^{\top},$$

Eigendecomposition

- di mana $\boldsymbol{Q}$ adalah matriks ortogonal yang terdiri dari vektor eigen dari $\boldsymbol{A}$, dan $\boldsymbol{\Lambda}$ adalah matriks diagonal.
- Nilai eigen $\boldsymbol{\Lambda}_{i,i}$ dikaitkan dengan vektor eigen di kolom *i* dari $\boldsymbol{Q}$, dilambangkan sebagai $\boldsymbol{Q}_{:,i}$. Karena $\boldsymbol{Q}$ adalah matriks ortogonal, kita dapat menganggap $\boldsymbol{A}$ sebagai ruang skala dengan λ_i dalam arah $\boldsymbol{v}^{(i)}$

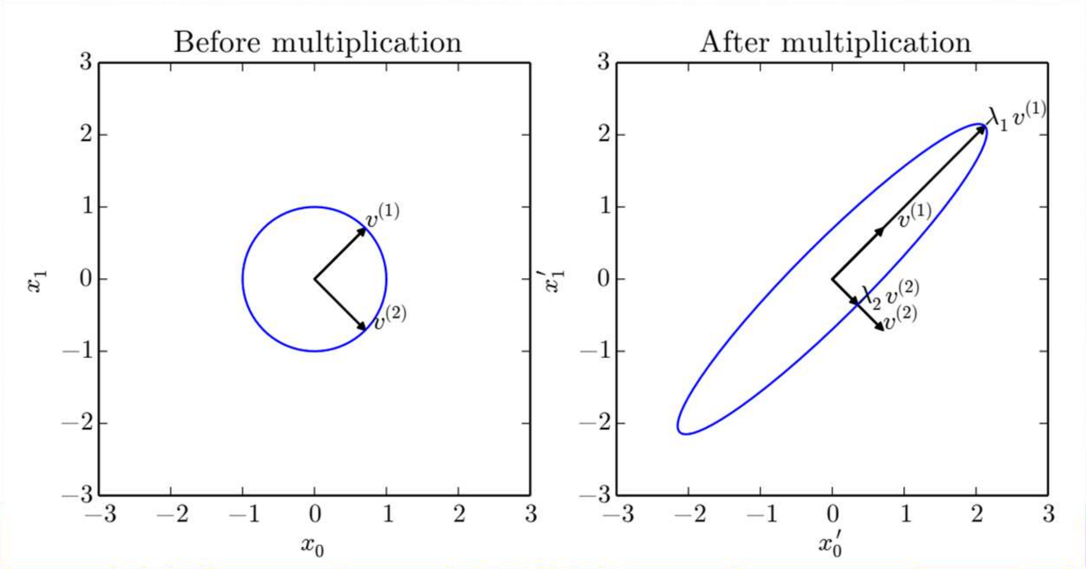
Eigendecomposition

- Two eigenvectors of a matrix



Contoh pengaruh vektor eigen dan nilai eigen

- Di sini, kita memiliki matriks ***A*** dengan dua vektor eigen ortonormal, $\boldsymbol{v}^{(1)}$ dengan nilai eigen λ_1 dan $\boldsymbol{v}^{(2)}$ dengan nilai eigen λ_2.
- (Kiri) Kita memplot himpunan semua vektor satuan $\boldsymbol{u} \in \mathbb{R}^2$ sebagai lingkaran satuan.
- (Kanan) Kita memplot himpunan semua titik ***Au***. Dengan mengamati cara ***A*** mendistorsi lingkaran satuan, kita dapat melihat bahwa ***A*** menskalakan ruang ke arah $\boldsymbol{v}^{(i)}$ dengan λ_i



Singular Value Decomposition

- Dekomposisi nilai singular serupa, kecuali kali ini kita akan menulis ***A*** sebagai hasil kali dari tiga matriks:

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\top.$$

- Misalkan ***A*** adalah matriks $m \times n$. Kemudian ***U*** didefinisikan menjadi matriks $m \times m$, ***D*** menjadi matriks $m \times n$, dan ***V*** menjadi matriks $n \times n$
- Lebih umum; matriks tidak harus persegi

Singular Value Decomposition

- Unsur-unsur <u>sepanjang diagonal</u> ***D*** dikenal sebagai <u>**nilai singular (singular value)**</u> matriks A.
- <u>Kolom</u> ***U*** dikenal sebagai <u>**vektor singular kiri (left-singular vector)**</u>.
- <u>Kolom</u> ***V*** dikenal sebagai <u>**vektor singular kanan (right-singular vector)**</u>.

Moore-Penrose Pseudoinverse

- Matrix inversion tidak bisa didefinisikan untuk matriks yang tidak persegi.
- Misalkan kita ingin membuat left-inverse ***B*** dari matriks ***A*** sedemikian hingga kita dapat menyelesaikan persamaan linier

$$\boldsymbol{Ax = y}$$

- dengan mengalikan kiri setiap sisi untuk mendapatkan

$$\boldsymbol{x = By}.$$

Moore-Penrose Pseudoinverse

- Pseudoinverse Moore-Penrose memungkinkan kita membuat beberapa kemajuan dalam kasus ini.
- Invers semu (pseudoinverse) dari ***A*** didefinisikan sebagai matriks

$$\boldsymbol{A}^{+} = \lim_{\alpha \searrow 0} (\boldsymbol{A}^{\top}\boldsymbol{A} + \alpha \boldsymbol{I})^{-1}\boldsymbol{A}^{\top}.$$

Moore-Penrose Pseudoinverse

- Algoritme praktis untuk menghitung pseudoinverse tidak didasarkan pada definisi ini, melainkan pada bentuknya:

$$A^{+} = V D^{+} U^{\top},$$

- di mana ***U***, ***D*** dan ***V*** adalah dekomposisi nilai singular dari ***A***, dan invers semu $\mathbf{D}^{+}$ dari matriks diagonal ***D*** diperoleh dengan mengambil kebalikan dari elemen bukan nolnya kemudian mengambil transpos dari matriks yang dihasilkan

Moore-Penrose Pseudoinverse

- Ketika ***A*** memiliki lebih banyak kolom daripada baris, maka menyelesaikan persamaan linier menggunakan invers semu memberikan **satu** dari banyak kemungkinan solusi.
- Secara khusus, ini memberikan solusi $\boldsymbol{x} = \boldsymbol{A}^{+} \boldsymbol{y}$ dengan norma Euclidean minimal $||\boldsymbol{x}||_2$ di antara semua kemungkinan solusi.
- Ketika ***A*** memiliki lebih banyak baris daripada kolom, mungkin saja **tidak ada solusi**. Dalam kasus ini, menggunakan pseudoinvers memberi kita ***x*** di mana ***Ax*** sedekat mungkin dengan ***y*** dalam hal norma Euclidean $||\boldsymbol{Ax} - \boldsymbol{y}||_2$.

Trace Operator

- Operator jejak (Trace Operator) memberikan jumlah dari semua entri diagonal dari sebuah matriks

$$\mathrm{Tr}(\boldsymbol{A}) = \sum_i \boldsymbol{A}_{i,i}.$$

- Jejak matriks persegi yang terdiri dari banyak faktor juga invarian untuk memindahkan faktor terakhir ke posisi pertama, jika bentuk matriks yang sesuai memungkinkan produk yang dihasilkan dapat ditentukan:

$$\mathrm{Tr}(\boldsymbol{ABC}) = \mathrm{Tr}(\boldsymbol{CAB}) = \mathrm{Tr}(\boldsymbol{BCA})$$

$$\mathrm{Tr}\Big(\prod_{i=1}^{n} \boldsymbol{F}^{(i)}\Big) = \mathrm{Tr}\Big(\boldsymbol{F}^{(n)} \prod_{i=1}^{n-1} \boldsymbol{F}^{(i)}\Big).$$

Trace Operator

- Contoh matriks berikut:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

- Maka diperoleh:
  - Tr{**A**} = 1+5+9 = 15 .

TERIMA KASIH