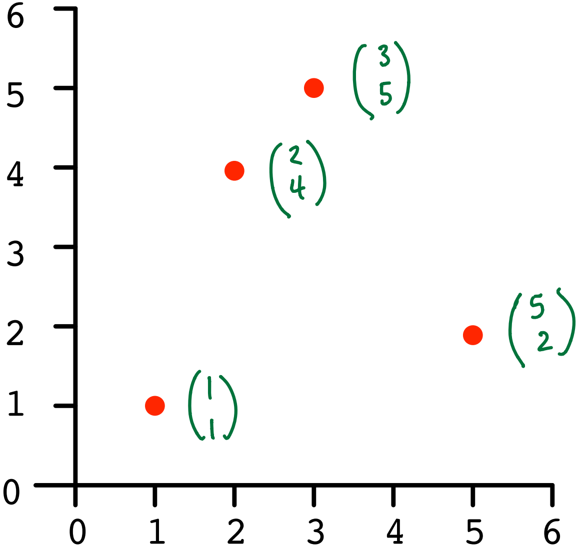
$Y = 2X$

$$\mathrm{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = \mathbb{E}[2X^2] - \mathbb{E}[X]\,\mathbb{E}[2X]$$

Linear algebra primer

DSE 210

Data as vectors and matrices



data set as a single matrix

$$\begin{bmatrix} 3 & 5 \\ 2 & 4 \\ 5 & 2 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 & 5 & 1 \\ 5 & 4 & 2 & 1 \end{bmatrix}$$

Matrix-vector notation

d features/coordinates, each real-valued

Vector $x \in \mathbb{R}^d$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_d \end{pmatrix} \quad d$$

d-dim vector

d×1

row 1, col 2

Matrix $M \in \mathbb{R}^{r \times d}$:

$$M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1d} \\ M_{21} & M_{22} & \cdots & M_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ M_{r1} & M_{r2} & \cdots & M_{rd} \end{pmatrix} \quad r$$

d

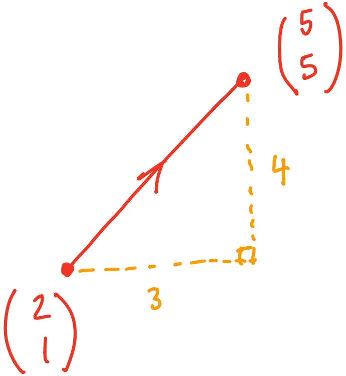
r×d

rows   #cols

M_{ij} = entry at row i, column j

Length of a vector

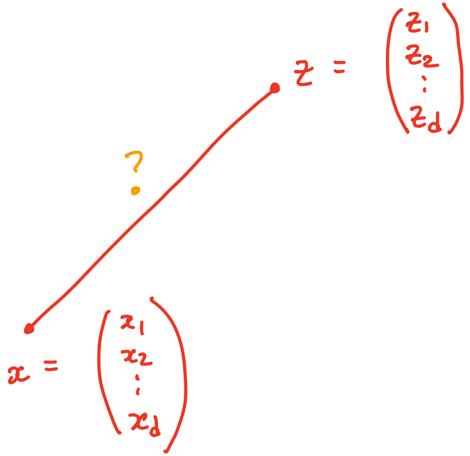
$\mathbb{R}^2$



$$\text{dist} = \sqrt{3^2 + 4^2} \quad \left(\begin{matrix}\text{Pythagorean} \\ \text{thm}\end{matrix}\right)$$

$$= 5$$

$\mathbb{R}^d$



distance between x and z

$$\|x - z\| = \sqrt{\sum_{i=1}^{d} (x_i - z_i)^2}$$

Length of a vector $x \in \mathbb{R}^d$ = distance from origin:

$$\|x\| = \sqrt{\sum_{i=1}^{d} x_i^2}$$

Q: What is the length of $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^d$?

$\sqrt{d}$

Transpose of vectors and matrices

$$\underset{4\times 1}{x} = \begin{pmatrix} 1 \\ 6 \\ 3 \\ 0 \end{pmatrix} \text{ has } \textbf{transpose} \quad \underset{1\times 4}{x^T} = \begin{pmatrix} 1 & 6 & 3 & 0 \end{pmatrix}$$

$$\underset{3\times 4}{M} = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 3 & 9 & 1 & 6 \\ 8 & 7 & 0 & 2 \end{pmatrix} \text{ has } \textbf{transpose} \quad \underset{4\times 3}{M^T} = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 9 & 7 \\ 0 & 1 & 0 \\ 4 & 6 & 2 \end{bmatrix}.$$

- $(A^T)_{ij} = A_{ji}$
- $(A^T)^T = A$

Adding and subtracting vectors and matrices

$$\begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 \\ 8 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -6 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 4 & 6 & 2 \\ 1 & 0 & 9 \end{pmatrix} + \begin{pmatrix} 0 & 8 & 1 \\ 3 & -10 & -9 \end{pmatrix} = \begin{pmatrix} 4 & 14 & 3 \\ 4 & -10 & 0 \end{pmatrix}$$

1. Dimensions <u>must</u> agree
2. Add / subtract elementwise

Dot product of two vectors

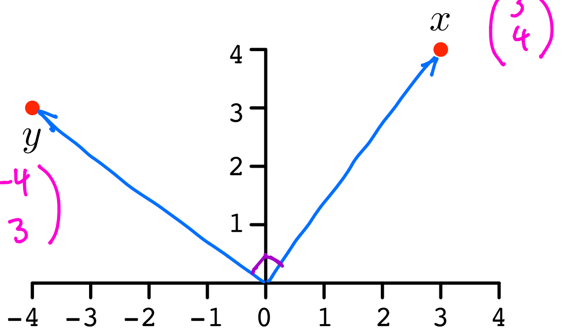
or "inner product"

Dot product of vectors $x, y \in \mathbb{R}^d$: dimensions must agree

$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$ $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{pmatrix}$

$$x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_d y_d.$$

What is the dot product between these two vectors?



$x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ $y = \begin{pmatrix} -4 \\ 3 \end{pmatrix}$

$$x \cdot y = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -4 \\ 3 \end{pmatrix}$$

$$= 3 \times (-4) + 4 \times 3 = 0$$

Dot products and angles

x orthogonal to y

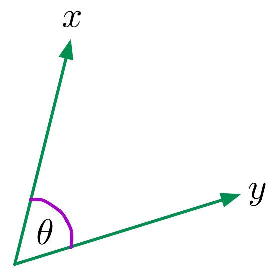
$\iff$ angle $= \pi/2$ (ie. 90°)

$\iff \cos(\text{angle}) = 0$

$\iff x \cdot y = 0$

Dot product of vectors $x, y \in \mathbb{R}^d$: $x \cdot y \;=\; x_1 y_1 + x_2 y_2 + \cdots + x_d y_d$.

Tells us the angle between x and y:



$$\cos\theta \;=\; \frac{x \cdot y}{\|x\|\,\|y\|}.$$

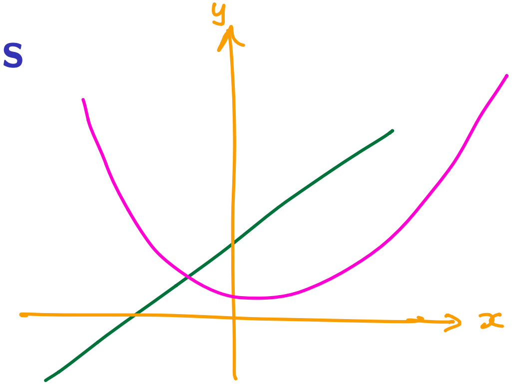
- x is **orthogonal** (at right angles) to y if and only if $x \cdot y = 0$
- When x, y are **unit vectors** (length 1): $\cos\theta = x \cdot y$
- What is $x \cdot x$?

$x \cdot x = x_1^2 + x_2^2 + \cdots + x_d^2 = \|x\|^2$

Another way to see it: angle btwn x, x is 0.

$\therefore x \cdot x = \|x\|\,\|x\| \cos 0 = \|x\|^2$

Linear and quadratic functions



In one dimension:

- Linear: $f(x) = 3x + 2$
- Quadratic: $f(x) = 4x^2 - 2x + 6$

Linear and quadratic functions

In one dimension:

- Linear: $f(x) = 3x + 2$
- Quadratic: $f(x) = 4x^2 - 2x + 6$

In higher dimension, e.g. $x = (x_1, x_2, x_3)$:

- Linear: $3x_1 - 2x_2 + x_3 + 4$
- Quadratic: $x_1^2 - 2x_1x_3 + 6x_2^2 + 7x_1 + 9$

Linear functions and dot products

$y = mx + b$

slope   y-intercept

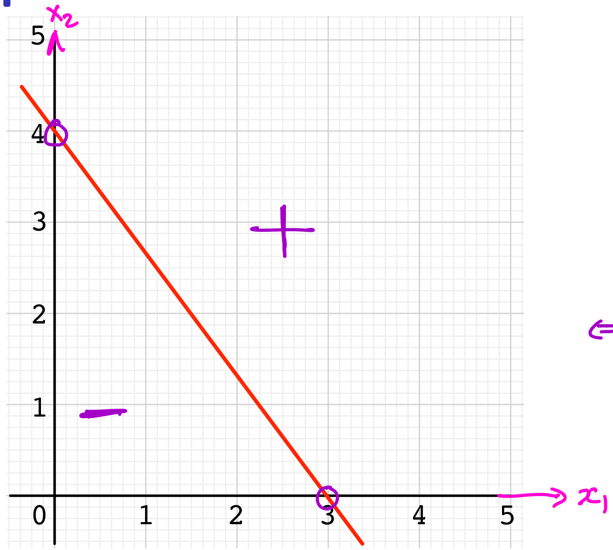
$m = -4/3$

$b = 4$

$x_2 = -\frac{4}{3}x_1 + 4$

$\iff 4x_1 + 3x_2 = 12$

Linear separator
$4x_1 + 3x_2 = 12$:



For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, linear separators are of the form:

$$w_1x_1 + w_2x_2 + \cdots + w_dx_d = c.$$

Can write as $w \cdot x = c$, for $w = (w_1, \ldots, w_d)$.

More general linear functions

A linear function from $\mathbb{R}^4$ to $\mathbb{R}$: $f\underbrace{(x_1, x_2, x_3, x_4)} = 3x_1 - 2x_3$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \qquad f(x) = w \cdot x \qquad \text{What is } w? \qquad w = \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}$$

A linear function from $\mathbb{R}^4$ to $\mathbb{R}^3$: $f(x_1, x_2, x_3, x_4) = (4x_1 - x_2, x_3, -x_1 + 6x_4)$

$$f(x) = \begin{pmatrix} 4x_1 - x_2 \\ x_3 \\ -x_1 + 6x_4 \end{pmatrix} = \underset{3 \times 4}{\begin{pmatrix} 4 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 6 \end{pmatrix}} \underset{4 \times 1}{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}} \in \underbrace{\mathbb{R}^3}_{\text{3-d vector}}$$

Matrix-vector product

Product of matrix $M \in \mathbb{R}^{r \times d}$ and vector $x \in \mathbb{R}^d$:

$$\underbrace{\begin{bmatrix} \text{———} \; M_1 \; \text{———} \\ \text{———} \; M_2 \; \text{———} \\ \vdots \\ \text{———} \; M_r \; \text{———} \end{bmatrix}}_{r \times d} \underbrace{\begin{bmatrix} | \\ x \\ | \end{bmatrix}}_{d \times 1} = \underbrace{\begin{bmatrix} M_1 \cdot x \\ M_2 \cdot x \\ \vdots \\ M_r \cdot x \end{bmatrix}}_{r \times 1}$$

inner dimensions must agree

take the dot product of each row of M with vector x

The identity matrix

The $d \times d$ **identity matrix** I_d sends each $x \in \mathbb{R}^d$ to itself.

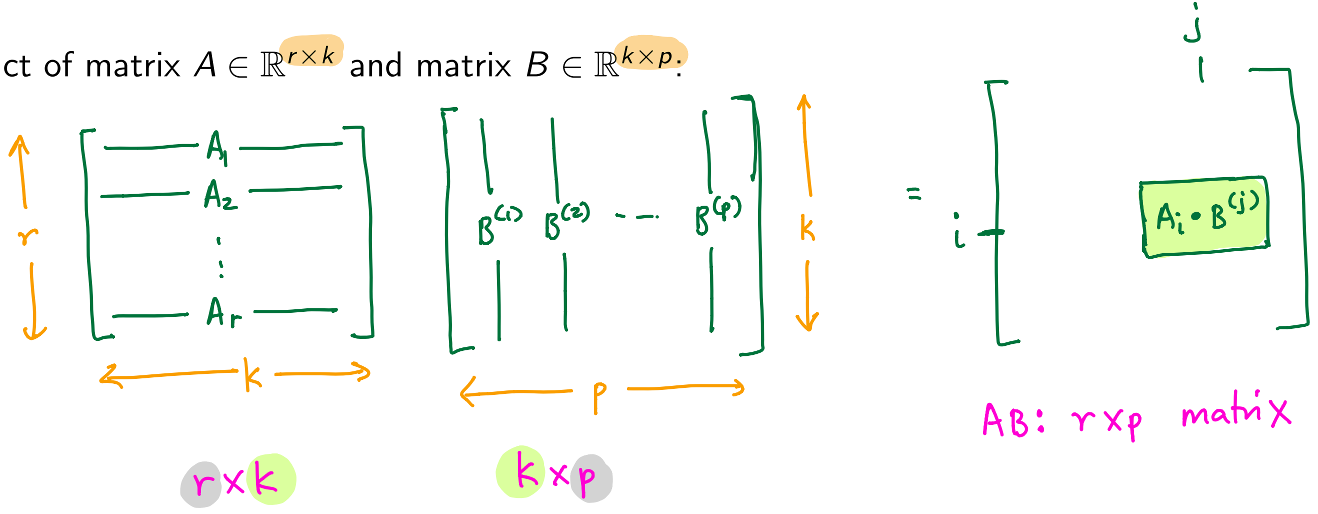
$$I_d = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

For any $x \in \mathbb{R}^d$,

$$I_d \, x = x$$

Matrix-matrix product

Product of matrix $A \in \mathbb{R}^{r \times k}$ and matrix $B \in \mathbb{R}^{k \times p}$.



AB is an $r \times p$ matrix whose (i,j) entry is

$$(AB)_{ij} = A_i \cdot B^{(j)} = \sum_{\ell=1}^{k} A_{i\ell} B_{\ell j}$$

Example:

$$\begin{bmatrix} 2 & 0 & 1 & 6 \\ 3 & 9 & 1 & 4 \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 17 & 6 & 14 \\ 34 & 13 & 11 \end{bmatrix}$$

2×4 $\quad$ 4×3 $\quad$ 2×3

Matrix products

If $A \in \mathbb{R}^{r \times k}$ and $B \in \mathbb{R}^{k \times p}$, then AB is an $r \times p$ matrix with (i,j) entry

$$(AB)_{ij} = (\text{dot product of } i\text{th row of } A \text{ and } j\text{th column of } B)$$
$$= \sum_{\ell=1}^{k} A_{i\ell} B_{\ell j}$$

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix}$$

$d \times 1$ $\quad$ $d \times 1$

- $I_k B = B$ and $A I_k = A$
- Can check: $(AB)^T = B^T A^T$
- For two vectors $u, v \in \mathbb{R}^d$, what is $u^T v$?

$$u^T v = \underbrace{\begin{pmatrix} u_1 & u_2 & \cdots & u_d \end{pmatrix}}_{1 \times d} \underbrace{\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix}}_{d \times 1} = u \cdot v$$

$u \cdot v \equiv u^T v$

when u, v are vectors of the same dimensions

Some special cases

$$\underset{d\times 1}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

For vector $x \in \mathbb{R}^d$, what are $x^T x$ and xx^T?

① $\underbrace{x^T x}_{1\times 1} = x \cdot x = \|x\|^2$

② $\underbrace{xx^T}_{d\times d}$ is a $d \times d$ matrix whose (i,j) entry is $x_i x_j$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} \begin{pmatrix} x_1 & x_2 & \cdots & x_d \end{pmatrix}$$

E.g. $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

$$xx^T = \begin{pmatrix} x_1^2 & x_1x_2 & x_1x_3 \\ x_2x_1 & x_2^2 & x_2x_3 \\ x_3x_1 & x_3x_2 & x_3^2 \end{pmatrix}$$

Associative but not commutative

- Multiplying matrices is **not commutative**: in general, $AB \neq BA$

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

$x^T(xx^T)(xx^T)(xx^T)x$

- But it is **associative**: $ABCD = (AB)(CD) = (A(BC))D$, etc.

Example: if $x \in \mathbb{R}^d$ has length 2, what is $x^T x x^T x x^T x x^T x$?

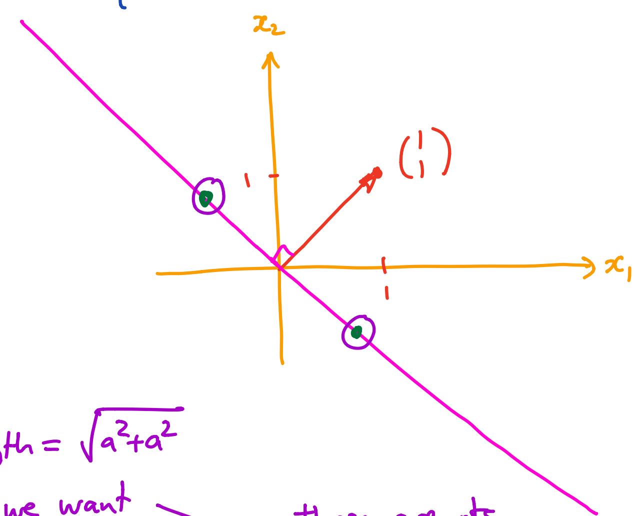
$$= (x^T x)(x^T x)(x^T x)(x^T x)$$

$$= \|x\|^2 \cdot \|x\|^2 \cdot \|x\|^2 \cdot \|x\|^2$$

$$= 4 \cdot 4 \cdot 4 \cdot 4 = 256$$

All points with $x \cdot x = 25$

$\{x \in \mathbb{R}^d : \|x\| = 5\}$



length $= \sqrt{a^2+a^2}$

if we want length $= 1$,

$2a^2 = 1 \Rightarrow a = \pm\frac{1}{\sqrt{2}}$

these are pts of the form $\begin{pmatrix}a\\-a\end{pmatrix}$

anything on this line is orthogonal to $\begin{pmatrix}1\\1\end{pmatrix}$

$\underbrace{\begin{pmatrix}1\\-1\end{pmatrix}}_{\text{length } \sqrt{2}}$ $\underbrace{\begin{pmatrix}-1\\1\end{pmatrix}}_{\text{length } \sqrt{2}}$

$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$ and $\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$

$(x_1, x_2) \cdot (1,1) = 0$

$\Leftrightarrow x_1 + x_2 = 0$

$\Leftrightarrow x_2 = -x_1$

A special case

$$x^T x = x^T \underset{\substack{d \times d \\ \text{identity}}}{I} x = x_1^2 + x_2^2 + \cdots + x_d^2$$

Recall: For vector $x \in \mathbb{R}^d$, we have $\boxed{x^T x = \|x\|^2.}$

What about $\underbrace{x^T}_{1 \times d} \underbrace{M}_{d \times d} \underbrace{x}_{d \times 1}$, for arbitrary $d \times d$ matrix M?

$$x^T M x = \sum_{i,j} M_{ij} x_i x_j$$

What is $x^T M x$ for $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$?

$$x^T M x = M_{11} x_1^2 + M_{12} x_1 x_2 + M_{21} x_2 x_1 + M_{22} x_2^2$$

$$= 1 \cdot x_1^2 + 2 x_1 x_2 + 0\, x_2 x_1 + 3 x_2^2$$

$$= \underbrace{x_1^2 + 2 x_1 x_2 + 3 x_2^2}_{\text{quadratic function of } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}$$

Mx linear fn
$x^T M x$ quadratic fn

What is $x^T M x$ for $M = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$?

$$\underset{1\times d}{x^T}\ \underset{d\times d}{M}\ \underset{d\times 1}{x} = (x_1\ x_2\ \cdots\ x_d) \underbrace{\begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1d} \\ M_{21} & M_{22} & & M_{2d} \\ \vdots & & & \vdots \\ M_{d1} & M_{d2} & \cdots & M_{dd} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}}_{Mx}$$

$$= \sum_{i=1}^{d} x_i (Mx)_i$$

$$= \sum_{i=1}^{d} x_i \sum_{j=1}^{d} M_{ij} x_j = \sum_{i,j=1}^{d} M_{ij} x_i x_j$$

Quadratic functions

Let M be any $d \times d$ (**square**) matrix.
For $x \in \mathbb{R}^d$, the mapping $x \mapsto x^T M x$ is a **quadratic function** from $\mathbb{R}^d$ to $\mathbb{R}$:

$$x^T M x = \sum_{i,j=1}^{d} M_{ij} x_i x_j.$$

What is the quadratic function associated with $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 3 & 4 & 5 \end{pmatrix}$?

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$x_1^2 + 2x_2^2 + 5x_3^2 + 3x_1x_3 + 4x_2x_3$$

Write the quadratic function $f(x_1, x_2) = \underbrace{x_1^2 + 2x_1x_2 + 3x_2^2}_{x^T M x}$ using matrices and vectors.

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\underset{2\times 2}{M} : \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} \checkmark$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \checkmark$$

$$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \checkmark$$

Special cases of square matrices

- **Symmetric**: $M = M^T$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix} \checkmark, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 3 & 4 & 6 \end{pmatrix} \times$$

- **Diagonal**: $M = \mathrm{diag}(m_1, m_2, \ldots, m_d)$

$$\underbrace{\mathrm{diag}(1, 4, 7)}_{\text{shorthand}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

Determinant of a square matrix

Determinant of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $|A| = ad - bc$.

Example: $A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \leftarrow |A| = 3 \times 2 - 1 \times 1 = 5$

The determinant is also defined for dxd matrices.

Determinant of $\begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_d \end{pmatrix}$ is $a_1 a_2 \cdots a_d$

$$|\mathrm{diag}(a_1, \ldots, a_d)| = a_1 a_2 \cdots a_d$$

Inverse of a square matrix

The **inverse** of a $d \times d$ matrix A is a $d \times d$ matrix B for which $AB = BA = I_d$.
Notation: A^{-1}.

Example: if $A = \begin{pmatrix} 1 & 2 \\ -2 & 0 \end{pmatrix}$ then $A^{-1} = \begin{pmatrix} 0 & -1/2 \\ 1/2 & 1/4 \end{pmatrix}$. Check!

$$\begin{pmatrix} 1 & 2 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1/2 \\ 1/2 & 1/4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Inverse of a square matrix, cont'd

The **inverse** of a $d \times d$ matrix A is a $d \times d$ matrix B for which $AB = BA = I_d$.
Notation: A^{-1}.

Square matrix A is singular if and only if $|A| = 0$

- Not all square matrices have an inverse
- Square matrix A is invertible if and only if $|A| \neq 0$
- What is the inverse of $A = \text{diag}(a_1, \ldots, a_d)$?

① Is $\text{diag}(a_1, .., a_d)$ necessarily invertible?

Not if one of a_i's is zero.

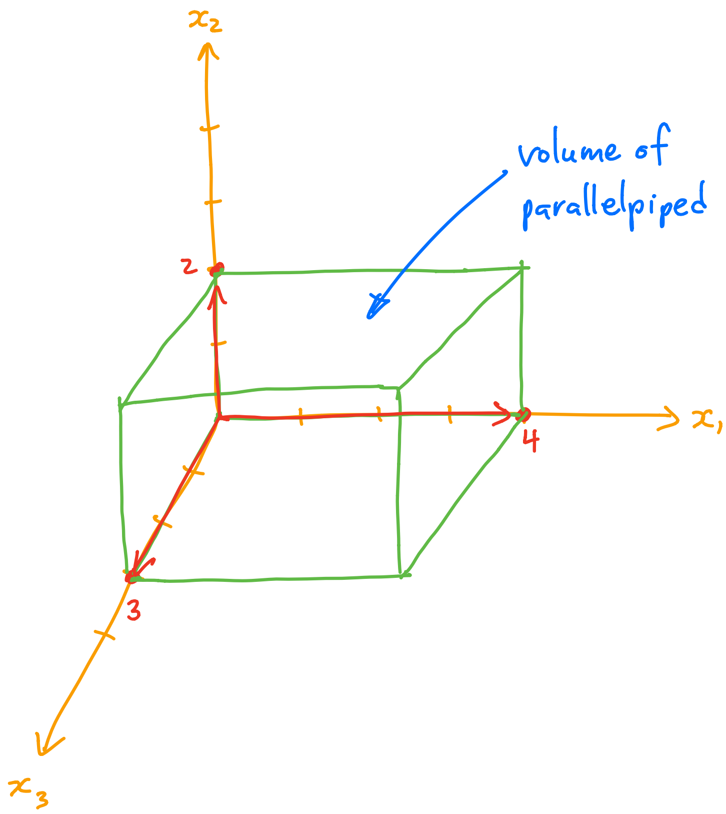
$\text{diag}(a_1, .., a_d)$ invertible if and only if all a_i are $\neq 0$

has inverse ≡ invertible

does not have inverse ≡ singular

② If $\text{diag}(a_1, .., a_d)$ is invertible, what is the inverse? It is $\text{diag}\left(\frac{1}{a_1}, \frac{1}{a_2}, .., \frac{1}{a_d}\right)$.

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$



volume of parallelpiped $= 2 \cdot 3 \cdot 4$
$= 24$