Some linear algebra background

CSE 250B

Matrix-vector notation

Vector $x \in \mathbb{R}^p$ and matrix $M \in \mathbb{R}^{r \times p}$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{pmatrix}, \quad M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1p} \\ M_{21} & M_{22} & \cdots & M_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ M_{r1} & M_{r2} & \cdots & M_{rp} \end{pmatrix}.$$

Matrix-vector notation

Vector $x \in \mathbb{R}^p$ and matrix $M \in \mathbb{R}^{r \times p}$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{pmatrix}, \quad M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1p} \\ M_{21} & M_{22} & \cdots & M_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ M_{r1} & M_{r2} & \cdots & M_{rp} \end{pmatrix}.$$

Transpose x^T and $M^T \in \mathbb{R}^{p \times r}$:

$$x^T = \begin{pmatrix} x_1 & x_2 & \cdots & x_p \end{pmatrix}, \quad M^T = \begin{pmatrix} M_{11} & \cdots & M_{r1} \\ M_{12} & \cdots & M_{r2} \\ M_{13} & \cdots & M_{r3} \\ \vdots & \ddots & \vdots \\ M_{1p} & \cdots & M_{rp} \end{pmatrix}.$$

Matrix-vector notation

Vector $x \in \mathbb{R}^p$ and matrix $M \in \mathbb{R}^{r \times p}$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{pmatrix}, \quad M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1p} \\ M_{21} & M_{22} & \cdots & M_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ M_{r1} & M_{r2} & \cdots & M_{rp} \end{pmatrix}.$$

Transpose x^T and $M^T \in \mathbb{R}^{p \times r}$:

$$x^T = \begin{pmatrix} x_1 & x_2 & \cdots & x_p \end{pmatrix}, \quad M^T = \begin{pmatrix} M_{11} & \cdots & M_{r1} \\ M_{12} & \cdots & M_{r2} \\ M_{13} & \cdots & M_{r3} \\ \vdots & \ddots & \vdots \\ M_{1p} & \cdots & M_{rp} \end{pmatrix}.$$

Properties of transpose: $(A^T)^T = A$ and $(AB)^T = B^T A^T$.

Dot product

Dot product of vectors $x, y \in \mathbb{R}^p$:

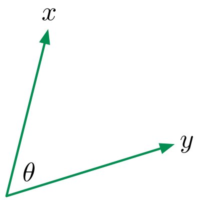
$$x \cdot y \;=\; x^T y \;=\; x_1 y_1 + \cdots + x_p y_p.$$

Dot product

Dot product of vectors $x, y \in \mathbb{R}^p$:

$$x \cdot y \;=\; x^T y \;=\; x_1 y_1 + \cdots + x_p y_p.$$

This tells us the angle between x and y:



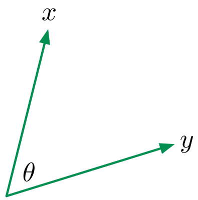
$$\cos\theta \;=\; \frac{x \cdot y}{\|x\| \, \|y\|}.$$

Dot product

Dot product of vectors $x, y \in \mathbb{R}^p$:

$$x \cdot y \;=\; x^T y \;=\; x_1 y_1 + \cdots + x_p y_p.$$

This tells us the angle between x and y:



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

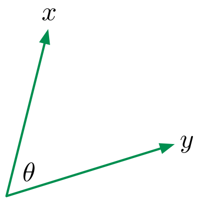
Easiest when x, y are *unit vectors* (length 1): then $\cos\theta = x \cdot y$.

Dot product

Dot product of vectors $x, y \in \mathbb{R}^p$:

$$x \cdot y \ = \ x^T y \ = \ x_1 y_1 + \cdots + x_p y_p.$$

This tells us the angle between x and y:



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

Easiest when x, y are *unit vectors* (length 1): then $\cos\theta = x \cdot y$.

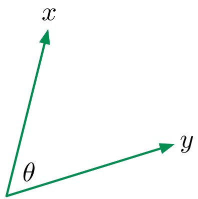
x is orthogonal (at right angles) to y iff $x \cdot y = ??$

Dot product

Dot product of vectors $x, y \in \mathbb{R}^p$:

$$x \cdot y \;=\; x^T y \;=\; x_1 y_1 + \cdots + x_p y_p.$$

This tells us the angle between x and y:



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

Easiest when x, y are *unit vectors* (length 1): then $\cos\theta = x \cdot y$.

x is orthogonal (at right angles) to y iff $x \cdot y = ??$

What is $x \cdot x$?

Matrix-vector products

If $M \in \mathbb{R}^{r \times p}$ and $x \in \mathbb{R}^p$ then

$$Mx = \begin{pmatrix} \longleftarrow & M_1 & \longrightarrow \\ \longleftarrow & M_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & M_r & \longrightarrow \end{pmatrix} \begin{pmatrix} \big\uparrow \\ x \\ \big\downarrow \end{pmatrix} = \begin{pmatrix} M_1 \cdot x \\ M_2 \cdot x \\ \vdots \\ M_r \cdot x \end{pmatrix}$$

Matrix-vector products

If $M \in \mathbb{R}^{r \times p}$ and $x \in \mathbb{R}^p$ then

$$Mx = \begin{pmatrix} \longleftarrow & M_1 & \longrightarrow \\ \longleftarrow & M_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & M_r & \longrightarrow \end{pmatrix} \begin{pmatrix} \uparrow \\ x \\ \downarrow \end{pmatrix} = \begin{pmatrix} M_1 \cdot x \\ M_2 \cdot x \\ \vdots \\ M_r \cdot x \end{pmatrix}$$

This mapping $x \mapsto Mx$ is a **linear function** from $\mathbb{R}^p$ to $\mathbb{R}^r$:

$$M(x + x') = Mx + Mx'.$$

Matrix-vector products

If $M \in \mathbb{R}^{r \times p}$ and $x \in \mathbb{R}^p$ then

$$Mx = \begin{pmatrix} \longleftarrow & M_1 & \longrightarrow \\ \longleftarrow & M_2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & M_r & \longrightarrow \end{pmatrix} \begin{pmatrix} \big\uparrow \\ x \\ \big\downarrow \end{pmatrix} = \begin{pmatrix} M_1 \cdot x \\ M_2 \cdot x \\ \vdots \\ M_r \cdot x \end{pmatrix}$$

This mapping $x \mapsto Mx$ is a **linear function** from $\mathbb{R}^p$ to $\mathbb{R}^r$:

$$M(x + x') = Mx + Mx'.$$

If $M \in \mathbb{R}^{p \times p}$ and $x \in \mathbb{R}^p$ then $x \mapsto x^T Mx$ is a **quadratic function** from $\mathbb{R}^p$ to $\mathbb{R}$:

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

Quick quiz

1. Write the linear function $f(x_1, x_2) = 3x_1 + 2x_2$ using vector notation (here, $x_1, x_2 \in \mathbb{R}$).

Quick quiz

1. Write the linear function $f(x_1, x_2) = 3x_1 + 2x_2$ using vector notation (here, $x_1, x_2 \in \mathbb{R}$).
2. Write the quadratic function $f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^2$ using matrices and vectors.

Quick quiz

1. Write the linear function $f(x_1, x_2) = 3x_1 + 2x_2$ using vector notation (here, $x_1, x_2 \in \mathbb{R}$).
2. Write the quadratic function $f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^2$ using matrices and vectors.
3. A linear function from $\mathbb{R}^3$ to $\mathbb{R}^3$ is given by the matrix

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

   As x varies, does Mx fill up all of $\mathbb{R}^3$?

A hierarchy of square matrices

Square $\qquad M \in \mathbb{R}^{p \times p}$

$\bigcup$

Symmetric $\qquad M = M^T$

$\bigcup$

Positive semidefinite $\qquad z^T M z \geq 0$ for all $z \in \mathbb{R}^p$

$\bigcup$

Positive definite $\qquad z^T M z > 0$ for all $z \neq 0$

Quick quiz

Symmetric matrix M is positive semidefinite (psd) if $z^T Mz \geq 0$ for all z.

1. PSD or not?

   - $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

Quick quiz

Symmetric matrix M is positive semidefinite (psd) if $z^T M z \geq 0$ for all z.

1. PSD or not?
   - $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
   - $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$

Quick quiz

Symmetric matrix M is positive semidefinite (psd) if $z^T M z \geq 0$ for all z.

1. PSD or not?
   - $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
   - $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$
2. A diagonal matrix is PSD if and only if ???

Quick quiz

Symmetric matrix M is positive semidefinite (psd) if $z^T M z \geq 0$ for all z.

1. PSD or not?

   - $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
   - $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$

2. A diagonal matrix is PSD if and only if ???
3. Show: If M, N are of the same size and PSD and $M + N$ is PSD.

Checking if a matrix is PSD

Useful fact: a matrix M is PSD iff it can be written in the form UU^T for some matrix U.

Checking if a matrix is PSD

Useful fact: a matrix M is PSD iff it can be written in the form UU^T for some matrix U.

Quick check: say $U \in \mathbb{R}^{r \times p}$ and $M = UU^T$.

Checking if a matrix is PSD

Useful fact: a matrix M is PSD iff it can be written in the form UU^T for some matrix U.

Quick check: say $U \in \mathbb{R}^{r \times p}$ and $M = UU^T$.

1. M is square.

Checking if a matrix is PSD

Useful fact: a matrix M is PSD iff it can be written in the form UU^T for some matrix U.

Quick check: say $U \in \mathbb{R}^{r \times p}$ and $M = UU^T$.

1. M is square.
2. M is symmetric.

Checking if a matrix is PSD

Useful fact: a matrix M is PSD iff it can be written in the form UU^T for some matrix U.

Quick check: say $U \in \mathbb{R}^{r \times p}$ and $M = UU^T$.

1. M is square.
2. M is symmetric.
3. Pick any $z \in \mathbb{R}^r$. Then

$$\begin{aligned} z^T M z \;=\; z^T U U^T z &\;=\; (z^T U)(U^T z) \\ &\;=\; (U^T z)^T (U^T z) \;=\; \|U^T z\|^2 \;\geq\; 0. \end{aligned}$$

Checking if a matrix is PSD

Useful fact: a matrix M is PSD iff it can be written in the form UU^T for some matrix U.

Quick check: say $U \in \mathbb{R}^{r \times p}$ and $M = UU^T$.

1. M is square.
2. M is symmetric.
3. Pick any $z \in \mathbb{R}^r$. Then

$$\begin{aligned} z^T M z \;=\; z^T U U^T z \;&=\; (z^T U)(U^T z) \\ &=\; (U^T z)^T (U^T z) \;=\; \|U^T z\|^2 \;\geq\; 0. \end{aligned}$$

Another useful fact: any covariance matrix is PSD. (Same argument, along with linearity of expectation.)

Eigenvalues and eigenvectors

1. Any matrix M defines a linear transformation $x \mapsto Mx$.

Eigenvalues and eigenvectors

1. Any matrix M defines a linear transformation $x \mapsto Mx$.
2. We'd like to understand the nature of these transformations. The easiest case is when M is **diagonal**:

$$\underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 2x_1 \\ -x_2 \\ 10x_3 \end{pmatrix}}_{Mx}$$

In this case, M simply scales each coordinate separately.

Eigenvalues and eigenvectors

1. Any matrix M defines a linear transformation $x \mapsto Mx$.
2. We'd like to understand the nature of these transformations. The easiest case is when M is **diagonal**:

$$\underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 2x_1 \\ -x_2 \\ 10x_3 \end{pmatrix}}_{Mx}$$

   In this case, M simply scales each coordinate separately.
3. What about more general matrices that are symmetric but not necessarily diagonal? They also just scale coordinates separately, but in a **different coordinate system**.

Eigenvalues and eigenvectors

1. Any matrix M defines a linear transformation $x \mapsto Mx$.
2. We'd like to understand the nature of these transformations. The easiest case is when M is **diagonal**:

$$\underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 2x_1 \\ -x_2 \\ 10x_3 \end{pmatrix}}_{Mx}$$

   In this case, M simply scales each coordinate separately.
3. What about more general matrices that are symmetric but not necessarily diagonal? They also just scale coordinates separately, but in a **different coordinate system**.

Let M be a $p \times p$ matrix.
We say $u \in \mathbb{R}^p$ is an **eigenvector** if M maps u onto the same direction, that is,

$$Mu = \lambda u$$

for some scaling constant λ. This λ is the **eigenvalue** associated with u.

Eigenvalues and eigenvectors: examples

We say u is an eigenvector of M, with eigenvalue λ, if $Mu = \lambda u$.

Eigenvalues and eigenvectors: examples

We say u is an eigenvector of M, with eigenvalue λ, if $Mu = \lambda u$.

Question: What are the eigenvectors and eigenvalues of:

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix} \text{ ?}$$

Eigenvalues and eigenvectors: examples

We say u is an eigenvector of M, with eigenvalue λ, if $Mu = \lambda u$.

Question: What are the eigenvectors and eigenvalues of:

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix} ?$$

Answer: Eigenvectors e_1, e_2, e_3, with corresponding eigenvalues $2, -1, 10$.

Eigenvalues and eigenvectors: examples

We say u is an eigenvector of M, with eigenvalue λ, if $Mu = \lambda u$.

Question: What are the eigenvectors and eigenvalues of:

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix} ?$$

Answer: Eigenvectors e_1, e_2, e_3, with corresponding eigenvalues $2, -1, 10$.

Question: Matrix $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ has eigenvectors

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

What are the corresponding eigenvalues?

Eigenvalues and eigenvectors: examples

We say u is an eigenvector of M, with eigenvalue λ, if $Mu = \lambda u$.

Question: What are the eigenvectors and eigenvalues of:

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix} \; ?$$

Answer: Eigenvectors e_1, e_2, e_3, with corresponding eigenvalues $2, -1, 10$.

Question: Matrix $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ has eigenvectors

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

What are the corresponding eigenvalues?
Answer: $\lambda_1 = 4$ and $\lambda_2 = 2$.

Eigenvalues and eigenvectors: examples

We say u is an eigenvector of M, with eigenvalue λ, if $Mu = \lambda u$.

Question: What are the eigenvectors and eigenvalues of:

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 10 \end{pmatrix} ?$$

Answer: Eigenvectors e_1, e_2, e_3, with corresponding eigenvalues $2, -1, 10$.

Question: Matrix $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ has eigenvectors

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

What are the corresponding eigenvalues?
Answer: $\lambda_1 = 4$ and $\lambda_2 = 2$.

In both cases the eigenvectors form an orthonormal basis.

Eigenvectors of a real symmetric matrix

Theorem. Let M be any real symmetric $p \times p$ matrix. Then M has

- p eigenvalues $\lambda_1, \ldots, \lambda_p$
- corresponding eigenvectors $u_1, \ldots, u_p \in \mathbb{R}^p$ that are **orthonormal**:

$$u_i \cdot u_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Eigenvectors of a real symmetric matrix

Theorem. Let M be any real symmetric $p \times p$ matrix. Then M has

- p eigenvalues $\lambda_1, \ldots, \lambda_p$
- corresponding eigenvectors $u_1, \ldots, u_p \in \mathbb{R}^p$ that are **orthonormal**:

$$u_i \cdot u_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

We can think of $u_1, \ldots, u_p$ as being the axes of the natural coordinate system for understanding M.

Eigenvectors of a real symmetric matrix

Theorem. Let M be any real symmetric $p \times p$ matrix. Then M has

- p eigenvalues $\lambda_1, \ldots, \lambda_p$
- corresponding eigenvectors $u_1, \ldots, u_p \in \mathbb{R}^p$ that are **orthonormal**:

$$u_i \cdot u_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

We can think of $u_1, \ldots, u_p$ as being the axes of the natural coordinate system for understanding M.

Theorem. Let M be any real symmetric $p \times p$ matrix, and let $\lambda_1, \ldots, \lambda_p$ be its eigenvalues. Then:

- M is positive semidefinite iff every λ_i is ≥ 0.
- M is positive definite iff every λ is > 0.