School of Computing
National University of Singapore
CS CS5240 Theoretical Foundations of Multimedia

Linear Algebra Review

We will use uppercase bold letters, $\mathbf{A}$, to denote matrices, lowercase bold letters, $\mathbf{x}$, to denote column vectors, and lowercase normal letters, a, to denote scalars. Thus:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \quad \text{where } a_{ij} \in \mathbb{R}$$

Above matrix has size $= m \times n$, i.e. m rows by n columns. If $m = n$, we say that $\mathbf{A}$ is square.

Vector: Column vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{x} = \mathbb{R}^n$

Row vector: $\mathbf{x}^\top = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$ where $^\top$ denotes the transpose operation.

Basic Operations

Transpose

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^\top = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Equality

$\mathbf{A} = \mathbf{B}$ iff same size and $a_{ij} = b_{ij}$ for all i,j

Addition, Multiplication

$k \in \mathbb{R}, \quad k\mathbf{A} = [ka_{ij}]$
$\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}], \qquad \mathbf{A}$,$\mathbf{B}$ same size
$\mathbf{C} = \mathbf{AB} \qquad \mathbf{A} : m \times r \quad \mathbf{B} : r \times n \quad \mathbf{C} : m \times n$
e.g. $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 1 & 0 \end{bmatrix}$
$c_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}$
Note:

$$\mathbf{AB} \neq \mathbf{BA}$$
$$(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$$

Matrix-Vector Multiplication

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

e.g. $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Linear Combination of columns...

$$(1)\begin{bmatrix} 1 \\ -1 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 0 \end{bmatrix} + (1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

In general,

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \ldots + x_n\mathbf{a}_n \qquad = \sum_{i=1}^{n} x_i\mathbf{a}_i$$

Similarly,

$$\begin{aligned} \mathbf{x}^\top \mathbf{A} &= \text{ row vector} \\ &= \text{ linear combination of rows of } \mathbf{A} \end{aligned}$$

Powers

$$\begin{aligned} &\mathbf{A}^2 = \mathbf{A}\mathbf{A} \\ &\mathbf{A}^k = \underbrace{\mathbf{A}\mathbf{A}\ldots\mathbf{A}}_{k} \\ &\mathbf{A}^0 = \mathbf{I} \qquad \text{(by convention)} \end{aligned}$$

Special Matrices

Zero

$\mathbf{0} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \\ 0 & 0 & \dots & 0 \end{bmatrix}$

Identity

$\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}$

$n \times n$ matrix

$1's$ along diagonal

$0's$ elsewhere

Triangular

e.g. $\begin{bmatrix}1 & 2 & 3\\ 0 & 2 & 1\\ 0 & 0 & 4\end{bmatrix}$ or $\begin{bmatrix}3 & 0 & 0\\ -1 & 1 & 0\\ 2 & 0 & 1\end{bmatrix}$

Diagonal

e.g. $\begin{bmatrix}5 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 6 & 0\\ 0 & 0 & 0 & -1\end{bmatrix}$

Symmetric

means $\quad \mathbf{A}^\top = \mathbf{A}$

e.g. $\begin{bmatrix}1 & 2 & 3\\ 2 & -1 & 4\\ 3 & 4 & 5\end{bmatrix}$

Skew-symmetric

means $\quad \mathbf{A}^\top = -\mathbf{A}$

Inverse $\mathbf{A}^{-1}$

$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$
Does not always exist. If $\mathbf{A}^{-1}$ exists, we say $\mathbf{A}$ is invertible or non-singular otherwise $\mathbf{A}$ is singular.
Note:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$\left(\mathbf{A}^\top\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\top$$

Proof:

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{A} &= \mathbf{I} \\ \Rightarrow \left(\mathbf{A}^{-1}\mathbf{A}\right)^\top &= \mathbf{I} \\ \Rightarrow \mathbf{A}^\top \underbrace{\left(\mathbf{A}^{-1}\right)^\top}_{(\mathbf{A}^\top)^{-1} \text{ by definition}} &= \mathbf{I} \end{aligned}$$

Thus we may write $\mathbf{A}^{-\top}$
Note: $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$
$\mathbf{AB} = \mathbf{0}$ does NOT mean $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$

Inner Product (Dot Product)

$$\begin{aligned}
\mathbf{x}^\top\mathbf{y} &= \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \\
&= \sum_{i=1}^{n} x_i y_i \qquad \text{(scalar)}
\end{aligned}$$

Outer Product

$$\begin{aligned}
\mathbf{x}\mathbf{y}^\top &= \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix} \\
&= \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ & & \vdots & \\ x_ny_1 & \dots & \dots & x_ny_n \end{bmatrix}
\end{aligned}$$

$\mathbf{x}\mathbf{y}^\top$ is singular (why?)

System of Linear Equations

Often, we need to solve:

$$\begin{aligned}
2x + y + z &= 5 \\
4x - 6y &= -2 \\
-2x + 7y + 2z &= 9
\end{aligned}$$

Rewrite in matrix form:

$$\begin{gathered}
\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} \\
\mathbf{A}\mathbf{x} = \mathbf{b}
\end{gathered}$$

Solution: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ if $\mathbf{A}^{-1}$ exists
Each equation represents a plane in 3D. Solution is the intersection of the planes. Possibilities:

(a) 3 planes parallel. No solution

(b) 2 planes parallel. No solution.

(c) No intersection. No solution.

(d) 3 planes coincident. Infinitely many solutions.

(e) 3 planes intersect in a line. Infinitely many solutions.

(f) 3 planes intersect at a point. Unique solution.

Later we will tackle the case when $\mathbf{A}$ is $m \times n$ (not square)

The 4 Fundamental Subspaces

Column space: $\text{Col}(\mathbf{A}) = \{$all possible linear combinations of cols. of $\mathbf{A}$ $\}$. Also known as $\text{Range}(\mathbf{A})$ or the span of the columns of $\mathbf{A}$.

$$\text{Let } \mathbf{A} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ \vdots & \vdots & & \vdots \end{bmatrix}. \quad \text{Then } \text{Col}(\mathbf{A}) = \left\{ \sum_i \lambda_i \mathbf{a}_i : \lambda_i \in \mathbb{R} \right\}$$

e.g.

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

$\text{Col}(\mathbf{A}) = \{$ all the vectors with 2nd component $= 0$ $\} = xz$ plane

Null space: $\text{Null}(\mathbf{A}) = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{0}\}$

For previous $\mathbf{A}$, $\text{Null}(\mathbf{A}) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$

$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ -1 & 3 & 2 \end{bmatrix}$

Note: 3rd col. of $\mathbf{B}$ = sum of 1st two cols.

$\text{Null}(\mathbf{B}) = \left\{ \lambda \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} : \lambda \in \mathbb{R} \right\}$

Note: $\text{Col}(\mathbf{B}) = \text{Col}(\mathbf{A})$, but nullspaces are different.

Similarly, for a matrix $\mathbf{A}$, we can define its *rowspace* as $\text{Col}(\mathbf{A}^\top)$; and its *left-nullspace* as $\text{Null}(\mathbf{A}^\top)$.

For a matrix $\mathbf{A}$, its rowspace is orthogonal to its nullspace, while its column space is orthogonal to its left-nullspace. Multiplication takes the rowspace of a matrix to its column space.

Linear Independence

A set of vectors $\mathbf{a}_1,\dots,\mathbf{a}_n$ is linearly independent if the only solution for

$$\sum_i \lambda_i \mathbf{a}_i = \mathbf{0}$$

is $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$

Linear dependence

$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \ldots + \lambda_n\mathbf{a}_n = \mathbf{0}$
say $\lambda_1 \neq 0$, then $\mathbf{a}_1 = -\frac{\lambda_2}{\lambda_1}\mathbf{a}_2 - \frac{\lambda_3}{\lambda_1}\mathbf{a}_3 \ldots - \frac{\lambda_n}{\lambda_1}\mathbf{a}_n$
i.e. we can express $\mathbf{a}_1$ as linear combination of $\mathbf{a}_2,\ldots,\mathbf{a}_n$
rank($\mathbf{A}$) = # linearly independent cols. of $\mathbf{A}$
nullity($\mathbf{A}$) = dimension of Null($\mathbf{A}$)
rank($\mathbf{A}$) + nullity($\mathbf{A}$) = # columns
Basically, rank counts the number of linear independent cols, nullity counts the number of linearly dependent cols.

Norm (length)

Euclidean or 2-norm: $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} = \sqrt{\mathbf{x}^\top\mathbf{x}}$
p-norm: $\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$
Euclidean distance between $\mathbf{x}$,$\mathbf{y}$: $\|\mathbf{x}-\mathbf{y}\|_2$
Cosine distance: $\mathbf{x}^\top\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta$
$-1 \leq \cos\theta = \frac{\mathbf{x}^\top\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1$
what's the difference between Euclidean and cosine distance?

Basis, Orthogonality

Consider 2D map, coordinate axes $\mathbf{i}$,$\mathbf{j}$
Any point $\mathbf{p}$ in 2D may be written as $\mathbf{p} = \alpha\mathbf{i} + \beta\mathbf{j}$ for some scalars α,β
$\mathbf{i}$,$\mathbf{j}$ are called *basis vectors.*
In fact, any 2 non-parallel vectors can be basis e.g. $\mathbf{p} = \alpha'\mathbf{a} + \beta'\mathbf{b}$
$\mathbf{i}$,$\mathbf{j}$ are "special" because they are orthonormal. i.e. unit length and 90° to each other.

Orthogonality

$\mathbf{x}$,$\mathbf{y}$ are orthogonal if $\mathbf{x}^\top\mathbf{y} = 0$

Orthonormal

A set of vectors $\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_k$ are orthonormal if
$\mathbf{b}_i^\top\mathbf{b}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$
In general, in $\mathbb{R}^n$, we need n vectors to form a basis. Prefer orthonormal basis because of convenience.
e.g. in 2D, $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ and $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$ form orthonormal basis.
Note: basis vectors are linearly independent., otherwise they cannot span (cover) the whole space.
A matrix $\mathbf{Q}$ is orthogonal if $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$, i.e. columns of $\mathbf{Q}$ are orthonormal.
Note: $\mathbf{Q}\mathbf{Q}^\top \neq \mathbf{I}$, unless $\mathbf{Q}$ is square.

e.g. $\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is an orthogonal matrix.

Gram-Schmidt Procedure

Input: $\mathbf{x}_1, \ldots, \mathbf{x}_n$ linearly independent vectors in $\mathbb{R}^n$
Output: $\mathbf{u}_1, \ldots, \mathbf{u}_n$ orthonormal basis.

1. Let $\mathbf{y}_1 = \mathbf{x}_1$, let $\mathbf{u}_1 = \dfrac{\mathbf{y}_1}{\|\mathbf{y}_1\|}$

2. $$\begin{aligned} \mathbf{y}_2 &= \mathbf{x}_2 - (\mathbf{x}_2^\top \mathbf{u}_1)\mathbf{u}_1 \\ \mathbf{u}_2 &= \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} \end{aligned}$$

   At every step :

   $$\begin{aligned} \mathbf{y}_k &= \mathbf{x}_k - \sum_{i=1}^{k-1} (\mathbf{x}_k^\top \mathbf{u}_i)\mathbf{u}_i \\ \mathbf{u}_k &= \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|} \end{aligned}$$

   Idea: Subtract away components that are represented in existing basis vectors.

e.g. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

1. $\mathbf{u}_1 = \dfrac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

2. $\mathbf{y}_2 = \mathbf{x}_2 - (\mathbf{x}_2^\top \mathbf{u}_1)\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \dfrac{1}{\sqrt{2}}(\dfrac{1}{\sqrt{2}})\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$

   $\mathbf{u}_2 = \dfrac{\mathbf{y}_2}{\|\mathbf{y}_2\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

3. $\mathbf{y}_3 = \mathbf{x}_3 - (\mathbf{x}_3^\top \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{x}_3^\top \mathbf{u}_2)\mathbf{u}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \sqrt{2}\left(\frac{1}{\sqrt{2}}\right)\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \sqrt{2}\left(\frac{1}{\sqrt{2}}\right)\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

   $\mathbf{u}_3 = \dfrac{\mathbf{y}_3}{\|\mathbf{y}_3\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Thus, $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

After Gram-Schmidt, $\mathbf{A} = \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\mathbf{Q}_{m\times n}} \underbrace{\begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \sqrt{2} \\ & \frac{1}{\sqrt{2}} & \sqrt{2} \\ & & 1 \end{bmatrix}}_{\mathbf{R}_{n\times n}}$

QR factorization!

In general, an $m \times n$ matrix $\mathbf{A}$ with linearly independent columns can be factored into $\mathbf{A} = \underbrace{\mathbf{Q}}_{m\times n} \underbrace{\mathbf{R}}_{n\times n}$

where, $\mathbf{Q}$ is orthogonal, $\mathbf{R}$ is upper triangular, square, invertible.

The columns of $\mathbf{Q}$ can be obtained using Gram-Schmidt on the cols. of $\mathbf{A}$.

The entries in $\mathbf{R}$ are the inner products computed during the Gram-Schmidt procedure.

$\mathbf{R} = \begin{bmatrix} \mathbf{x}_1^\top \mathbf{u}_1 & \mathbf{x}_2^\top \mathbf{u}_1 & \mathbf{x}_3^\top \mathbf{u}_1 \\ & \mathbf{x}_2^\top \mathbf{u}_2 & \mathbf{x}_3^\top \mathbf{u}_2 \\ & & \mathbf{x}_3^\top \mathbf{u}_3 \end{bmatrix}$

Alternatively, once we have $\mathbf{Q}$, then $\mathbf{R} = \mathbf{Q}^\top \mathbf{A}$.

Or, if we have $\mathbf{R}$, $\mathbf{Q} = \mathbf{A}\mathbf{R}^{-1}$. Note that $Col(\mathbf{Q}) = Col(\mathbf{A})$.

Eigenvalues/ Eigenvectors

For a square matrix $\mathbf{A}$, we often need to solve for $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x} \qquad \text{Eigenvector/eigenvalue problem.}$$

Where, $\mathbf{x}$ is the eigenvector. λ is a scalar (eigenvalue).

In general, a square matrix rotates and scales $\mathbf{x}$. But if $\mathbf{x}$ is an eigenvector, then $\mathbf{A}$ simply scales it (no rotation)

How to compute $\mathbf{x}$, λ? One way (only for small matrices) is to solve the nth degree polynomial:

$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$

Determinant, Trace

$\det(\mathbf{A})$ measures "size" of $\mathbf{A}$

For 2×2 $\quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det(\mathbf{A}) = ad - bc$

For triangular matrix, det = product of diagonal elements.

In general, $\det(\mathbf{A})$ follows a recursive formula. Slow to compute. So we use other tricks:

$\det(\mathbf{A})$ = product of eigenvalues $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$

$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$

$\det(\mathbf{A}) = 0 \iff \mathbf{A}$ singular

$\det(\mathbf{A}) = \det(\mathbf{A}^\top)$

$\det(k\mathbf{A}) = k^n \det(\mathbf{A}), \quad \mathbf{A}$ is $n \times n$

Trace

$$\begin{aligned}\operatorname{tr}(\mathbf{A}) &= \text{sum of diagonal elements}\\ &= \text{sum of eigenvalues}\end{aligned}$$

$$\begin{aligned}\operatorname{tr}(\mathbf{AB}) &= \operatorname{tr}(\mathbf{BA})\\ \operatorname{tr}(\mathbf{A}+\mathbf{B}) &= \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B})\end{aligned}$$

$\operatorname{tr}(\mathbf{A})$ also measures "size" of $\mathbf{A}$.

Back to eigenvector/eigenvalue

e.g. $\mathbf{A} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$

$$\begin{aligned}\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \qquad&\\ \Rightarrow \det\left(\begin{bmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{bmatrix}\right) &= 0\\ \Rightarrow (-5-\lambda)(-2-\lambda) - 4 &= 0\\ \Rightarrow \lambda^2 + 7\lambda + 6 &= 0\\ \Rightarrow (\lambda+1)(\lambda+6) &= 0\end{aligned}$$

roots: $\lambda_1 = -1$, $\lambda_2 = -6$
2 eigenvalues
Find $\mathbf{x}_1$: $\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} -\alpha \\ -\beta \end{bmatrix}$
since $\mathbf{Ax} = \lambda\mathbf{x}$
$\Rightarrow \quad 2\alpha = \beta$, so $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, corresponding to $\lambda_1 = -1$
Find $\mathbf{x}_2$: $\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} -6\alpha \\ -6\beta \end{bmatrix} \Rightarrow \quad \alpha = -2\beta$
so $\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, corresponding to $\lambda_2 = -6$ There are 2 eigenvectors.
In general, $\mathbf{A}$ $\quad n \times n$ has n eigenvalues and n eigenvectors. Note: They can be complex!
Note:

$$(k\mathbf{A})\mathbf{x} = (k\lambda)\mathbf{x} \tag{1}$$

$$\mathbf{A}(k\mathbf{x}) = \lambda(k\mathbf{x}) \tag{2}$$

$$\mathbf{A}^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x} \tag{3}$$

Because of Equation (2) above, from now on we will assume that an eigenvector has norm = 1. Since if it is not, we can simply divide it by its norm.

In matrix form: Let $\mathbf{E} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \vdots & \vdots & & \vdots \end{bmatrix}$ eigenvector matrix

$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$ diagonal eigenvalue matrix.

Then $\mathbf{AE} = \mathbf{E}\Lambda$

If $\mathbf{A}$ symmetric, then $\mathbf{E}$ is orthogonal and Λ is real, thus $\mathbf{A} = \mathbf{E}\Lambda\mathbf{E}^\top$ or $\mathbf{E}^\top\mathbf{AE} = \Lambda$. This is called the *Spectral Theorem.*

We say that $\mathbf{E}$ diagonalizes $\mathbf{A}$

Note:

$$\begin{aligned} \mathbf{A}^2 &= (\mathbf{E}\Lambda\mathbf{E}^\top)(\mathbf{E}\Lambda\mathbf{E}^\top) \\ &= \mathbf{E}\Lambda^2\mathbf{E}^\top \qquad \text{since } \mathbf{E}^\top\mathbf{E} = \mathbf{I} \\ \mathbf{A}^k &= (\mathbf{E}\Lambda\mathbf{E}^\top)(\mathbf{E}\Lambda\mathbf{E}^\top)\dots(\mathbf{E}\Lambda\mathbf{E}^\top) \\ &= \mathbf{E}\Lambda^k\mathbf{E} \\ \mathbf{A}^{-1} &= \mathbf{E}\Lambda^{-1}\mathbf{E}^\top \end{aligned}$$

Inverse of diagonal matrix: $\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & & & \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_n} \end{bmatrix}$

Cross product

This is defined only for vectors in $\mathbb{R}^3$. Let $\mathbf{a} = [a_1 \quad a_2 \quad a_3]^\top$, and $\mathbf{b} = [b_1 \quad b_2 \quad b_3]^\top$. Then the vector cross product is defined as:

$$\begin{aligned} \mathbf{c} &= \mathbf{a} \times \mathbf{b} \\ &= \det\left(\begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\right) \\ &= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \\ &= [(a_2b_3 - a_3b_2) \quad (a_1b_3 - a_3b_1) \quad (a_1b_2 - a_2b_1)]^\top \end{aligned}$$

Geometrically, $\mathbf{c}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. Norm: $||\mathbf{c}|| = ||\mathbf{a}||||\mathbf{b}||sin\theta$, where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$. The direction of $\mathbf{c}$ is usually determined by the *right-hand rule*: position the 4 fingers of your right hand over $\mathbf{a}$, and rotate around the thumb towards $\mathbf{b}$; the thumb points in the direction of $\mathbf{c}$. Note: some authors write $\mathbf{a} \wedge \mathbf{b}$ to denote cross product. Useful identities: for any 3 vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$,

$$\begin{aligned} \mathbf{a}^\top\mathbf{b} \times \mathbf{c} = \mathbf{b}^\top\mathbf{c} \times \mathbf{a} = \mathbf{c}^\top\mathbf{a} \times \mathbf{b} \\ \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a}^\top\mathbf{c})\mathbf{b} - (\mathbf{a}^\top\mathbf{b})\mathbf{c} \end{aligned}$$