Recurrence / Recursive relations: involve sequences $x_0, x_1, x_2, \dots$ for which each term x_n depends on the previous terms. The standard example: $1, 1, 2, 3, 5, 8, \dots$ Fibonacci sequence $x_n = x_{n-1} + x_{n-2}$. Knowing x_0, x_1 gives x_2. Knowing x_1, x_2 gives x_3. We need to find an efficient method for computing x_n.

Example 1. Variables / Populations depending on one another.

$$\left\{ \begin{aligned} x_{n+1} &= 1.02\, x_n - 0.25\, y_n \\ y_{n+1} &= 0.8\, x_n + 1.04\, y_n \end{aligned} \right\} \quad \circledast$$

Knowing x_1, y_1 gives x_2, y_2. Knowing x_n, y_n gives x_{n+1}, y_{n+1}.
In biology, x_n, y_n could be numbers of chickens, foxes during ~~day~~ day n.
Evolution of $x_n \to$ positive effect from x_n, negative effect from y_n
presence of x_n is helpful, presence of y_n is hurtful.
Evolution of $y_n \to$ presence of x_n, y_n is helpful.

Solution. To solve system $\circledast$, we need to determine $\vec{u}_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix}$.
The system says $\vec{u}_{n+1} = \begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 1.02\, x_n - 0.25\, y_n \\ 0.8\, x_n + 1.04\, y_n \end{bmatrix}$,

namely $\vec{u}_{n+1} = \begin{bmatrix} 1.02 & -0.25 \\ 0.8 & 1.04 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} = A \cdot \vec{u}_n$.

This gives

$$\vec{u}_1 = A \cdot \vec{u}_0$$

$$\vec{u}_2 = A \cdot \vec{u}_1 = A^2 \cdot \vec{u}_0$$

$$\vec{u}_3 = A\, \vec{u}_2 = A^3 \cdot \vec{u}_0$$

and $\vec{u}_n = A^n \vec{u}_0$ more generally. We need an efficient method for computing powers of the matrix A.

Example 2. We count the number of n-digit integers that contain the digits 1, 2, 3 but no consecutive ones?

Let x_n be the number of such integers. Look at first digit.

$$\underline{1}\ \underline{2}\ \underbrace{\_\ \_\ \_\ \_\ \_}_{x_{n-2} \text{ choices}} \qquad\qquad \underline{1}\ \underline{3}\ \underbrace{\_\ \_\ \_\ \_\ \_}_{x_{n-2} \text{ choices}}$$

$$\underline{2}\ \underbrace{\_\ \_\ \_\ \_\ \_}_{x_{n-1} \text{ choices}}$$

$$\underline{3}\ \underbrace{\_\ \_\ \_\ \_\ \_}_{x_{n-1} \text{ choices}}$$

This implies the relation

$$\boxed{x_n = 2x_{n-1} + 2x_{n-2}}$$

We get a recursive relation. Once again, we need x_1, x_2 to get x_3.

Define $\vec{u}_n = \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix}$, vector of two consecutive terms.

Then $$\vec{u}_{n+1} = \begin{bmatrix} x_{n+1} \\ x_{n+2} \end{bmatrix} = \begin{bmatrix} x_{n+1} \\ 2x_{n+1} + 2x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix}$$

So $\vec{u}_{n+1} = A \cdot \vec{u}_n$ as before.

<u>Powers of diagonal matrices.</u> Suppose $A = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$ is diagonal. An example is $A = \begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix}$ and $\begin{bmatrix} d_1 & \\ & d_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \end{bmatrix}$.
Left mult. by A sends x_i to $d_i x_i$. This suggests

$$A^k = \begin{bmatrix} d_1^k & \\ & d_2^k \end{bmatrix}$$ and we can prove this by induction.

<u>Theorem</u> If A is diagonal with diagonal entries $d_1, d_2, \dots, d_n$ then A^k is diagonal with diagonal entries $d_1^k, d_2^k, \dots, d_n^k$.

<u>Proof.</u> Since $A = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$, we have $A\vec{e}_i = d_i \vec{e}_i$ for each i.
We need to show $A^k = \begin{bmatrix} d_1^k & & \\ & \ddots & \\ & & d_n^k \end{bmatrix}$; namely $A^k \vec{e}_i = d_i^k \vec{e}_i$.
By induction. True if $k=1$. Assume for some k. Then

$$A^{k+1}\vec{e}_i = A \cdot A^k \vec{e}_i = A \cdot (d_i^k \vec{e}_i) = d_i^k (A\vec{e}_i) = d_i^{k+1}\vec{e}_i. \quad \blacksquare$$

<u>Powers of diagonalisable matrices.</u> Suppose A is a square matrix such that $B^{-1}AB = D$ is diagonal. In that case, we say A is diagonalisable. We can compute D^k. To compute A^k, note that

$$D^k = (B^{-1}AB)^k = B^{-1}AB \cdot B^{-1}AB \cdot B^{-1}AB \cdot \ldots \cdot B^{-1}AB$$
$$\Longrightarrow \quad D^k = B^{-1}A^k B.$$

This proves the formula

$$\boxed{A^k = B \cdot D^k \cdot B^{-1}.}$$

When is a matrix A diagonalisable?
Suppose $B^{-1}AB = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix}$ is diagonal and $B = [\vec{v}_1\ \vec{v}_2 \ldots \vec{v}_n]$.
Then

$$B^{-1}AB\vec{e}_i = d_i \vec{e}_i \text{ for each } i \iff AB\vec{e}_i = d_i B\vec{e}_i$$
$$\iff A\vec{v}_i = d_i \vec{v}_i$$

Definition We say $\vec{v}$ is an eigenvector of A with eigenvalue λ, if $\vec{v}$ is nonzero and $\boxed{A\vec{v} = \lambda\vec{v}}$.

To diagonalise A, if possible, we need to find the eigenvectors $v_1, v_2, \dots, v_n$ and merge them into a matrix B to get $B^{-1}AB$ diagonal.

Step 1. | Finding the eigenvalues λ.

We need to solve $A\vec{v} = \lambda\vec{v}$ or $(A-\lambda I)\vec{v} = 0$, where $\vec{v}$ is nonzero. We thus need $(A-\lambda I)\vec{v} = 0$ to not have a unique solution. If $A-\lambda I$ is invertible, then $\vec{v}=0$ is the only solution. We thus need $A-\lambda I$ to not be invertible (so the reduced row echelon form will contain free variables). We get eigenvectors if and only if $A-\lambda I$ is not invertible, hence if and only if $\boxed{\det(A-\lambda I) = 0}$. We call

$$f(\lambda) = \det(A-\lambda I) = \text{the characteristic polynomial of } A.$$

Step 2. | Finding the eigenvectors $\vec{v}$

Once λ is known, we can simply solve $A\vec{v} = \lambda\vec{v}$.

Example. Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$. To find the eigenvalues λ, we solve $\det(A-\lambda I) = 0$. We get

$$\begin{aligned}
\det(A-\lambda I) = \det \begin{bmatrix} 1-\lambda & 3 \\ 4 & 2-\lambda \end{bmatrix} &= (1-\lambda)(2-\lambda) - 12 \\
&= 2 - \lambda - 2\lambda + \lambda^2 - 12 \\
&= \lambda^2 - 3\lambda - 10 \\
&= (\lambda-5)(\lambda+2)
\end{aligned}$$

and the eigenvalues are $\lambda = 5$ and $\lambda = -2$. To find the eigenvectors, we solve $A\vec{v} = \lambda\vec{v}$ for each λ separately.

$$A\vec{v} = 5\vec{v}$$

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

$x + 3y = 5x$, $4x + 2y = 5y$

$3y = 4x$, $\cancel{3y = 4x}$

$y = \frac{4x}{3}$ and $\vec{v} = \begin{bmatrix} x \\ \frac{4x}{3} \end{bmatrix}$

Eigenvectors $\rightsquigarrow$ $v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, for instance
(or any scalar multiple)

$$A\vec{v} = -2\vec{v}$$

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2x \\ -2y \end{bmatrix}$$

$x + 3y = -2x$, $4x + 2y = -2y$

$y = -x$, $\cancel{y = -x}$

$y = -x$ and $\vec{v} = \begin{bmatrix} x \\ -x \end{bmatrix}$

Eigenvectors $\rightsquigarrow$ $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, for instance.

<u>Theorem (Diagonalisable matrices)</u> Suppose A is $n\times n$ and $B = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_n]$.
Then $B^{-1}AB = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots\ \lambda_n \end{bmatrix}$ is diagonal $\iff$ $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are lin. indep. with $\boxed{A\vec{v}_i = \lambda_i \vec{v}_i}$ for each i.
In particular, A is diagonalisable $\iff$ A has n lin. indep. eigenvectors.

<u>Proof.</u> We have $B\vec{e}_i = \vec{v}_i$ for all i. To say $B^{-1}AB = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots\ \lambda_n \end{bmatrix}$ is to say $B^{-1}AB\,\vec{e}_i = \lambda_i \vec{e}_i \iff AB\vec{e}_i = \lambda_i B\vec{e}_i \iff A\vec{v}_i = \lambda_i \vec{v}_i$.
Moreover, B is invertible if and only if $\vec{v}_1, \dots \vec{v}_n$ are lin. indep. ▨

<u>Note</u>: We need to find eigenvalues λ_i and eigenvectors $\vec{v}_i$.
Since $A\vec{v}_i = \lambda_i \vec{v}_i$ gives $A\vec{v}_i - \lambda_i \vec{v}_i = 0$ or $(A - \lambda I)\vec{v}_i = 0$,
eigenvalues λ_i are ----- roots of $f(\lambda) = \det(A - \lambda I)$ and
eigenvectors $\vec{v}_i$ are ----- solutions of $A\vec{v}_i = \lambda\vec{v}_i$ or $(A - \lambda I)\vec{v}_i = 0$.

<u>Theorem</u> (The 2×2 case) If A is a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$f(\lambda) = \det(A - \lambda I) = \lambda^2 - (a+d)\lambda + ad - bc$$
$$= \lambda^2 - \underbrace{(\operatorname{tr} A)}_{\text{trace of } A}\lambda + \det A.$$

Proof. $f(\lambda) = \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = (a-\lambda)(d-\lambda) - bc = ad - a\lambda - d\lambda + \lambda^2 - bc$

$= \lambda^2 - (a+d)\lambda + ad - bc.$ ∎

Example 1 (A diagonalisable matrix) Let $A = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$.

Eigenvalues $\lambda \longrightarrow$ roots of $f(\lambda) = \det(A - \lambda I) = \lambda^2 - 5\lambda + 6 = (\lambda-2)(\lambda-3)$ and thus $\lambda_1 = 2$ and $\lambda_2 = 3$.

Eigenvectors $\vec{v} \longrightarrow$ solutions of $(A - \lambda I)\vec{v} = 0$.

$\lambda = 2$

$(A - 2I)\vec{v} = 0$

$$\begin{bmatrix} 2 & -2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\cancel{2x - 2y = 0}$, $x - y = 0$

So $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ x \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

We get scalar multiples of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$\lambda = 3$

$(A - 3I)\vec{v} = 0$

$$\begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$x - 2y = 0$, $\cancel{x - 2y = 0}$

So $\vec{v} = \begin{bmatrix} 2y \\ y \end{bmatrix} = y \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

We get scalar multiples of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Check $B = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ is invertible with inverse $B^{-1} = -\begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$

and $B^{-1}AB = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$.

$$= \begin{bmatrix} -2 & 4 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \overset{\lambda_1}{2} & 0 \\ 0 & \overset{\lambda_2}{3} \end{bmatrix}$$

Similarly, $B = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow B^{-1}AB = \begin{bmatrix} 3 & \\ & 2 \end{bmatrix}$

and $B = \begin{bmatrix} 4 & 3 \\ 2 & 3 \end{bmatrix} \Rightarrow B^{-1}AB = \begin{bmatrix} 3 & \\ & 2 \end{bmatrix}$.

Theorem (Upper/lower triangular matrices) If A is upper/lower triangular, say $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & \vdots \\ 0 & & \ddots & \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}$, then the eigenvalues of A are the diagonal entries of A.

Proof. $f(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} a_{11}-\lambda & * & \cdots & * \\ & a_{22}-\lambda & \ddots & \vdots \\ & & \ddots & * \\ & & & a_{nn}-\lambda \end{bmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda).$ ∎

Example 2. (A non-diagonalisable matrix) Let $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$.

Then eigenvalues $\lambda \longrightarrow$ upper triangular so $\lambda = 2, 2$.

Eigenvectors $\vec{v} \longrightarrow$ we solve $(A-\lambda I)\vec{v} = 0$ or $(A-2I)\vec{v} = 0$.

In this case, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ so $y = 0$

and $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We get scalar multiples of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Thus, A does <u>not</u> have 2 lin. indep. eigenvectors and A is <u>not</u> diagonalisable. We get one eigenvalue and one eigenvector essentially.

Null space of a matrix

Eigenvectors $\vec{v}$ satisfy $(A-\lambda I)\vec{v} = 0$ so $\vec{v}$ belongs to the null space of $A-\lambda I$. There is a systematic method for finding the null space of a matrix.

Step 1. Use row reduction to find the reduced row echelon form. If the matrix is A and $R = E_k \cdots E_2 E_1 A$ is the RREF, then $A\vec{v} = 0 \iff E_k \cdots E_2 E_1 A \vec{v} = 0 \iff R\vec{v} = 0$.

This shows that A, R have the same null space.

Step 2. We compute the null space of R by eliminating the pivot variables.

Null space of a matrix

Recall that $N(A) = \{\text{vectors } \vec{v} \text{ such that } A \cdot \vec{v} = 0\}$. This is related to eigenvectors ... $A\vec{v} = \lambda \vec{v}$ or $(A - \lambda I)\vec{v} = 0$.
Thus, eigenvectors are <u>nonzero</u> elements of $N(A - \lambda I)$.

To find the null space of a matrix A, we first

① compute the RREF of the matrix noting that $N(A) = N(R)$.
This is because $R = E_k \cdots E_1 A$ implies

$$R\vec{v} = E_k \cdots E_1 A\vec{v}, \text{ so } A\vec{v} = 0 \iff R\vec{v} = 0.$$

② find a basis for $N(R)$ by eliminating the pivot variables.

Suppose, for instance, that $R = \begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

The null space of R consists of vectors $\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$ with

$$\begin{cases} \underline{x_1} + 2x_2 + 3x_4 = 0 \\ \underline{x_3} + x_4 = 0 \\ \underline{x_5} = 0 \end{cases}. \text{ This gives } \begin{cases} x_1 = -2x_2 - 3x_4 \\ x_3 = -x_4 \\ x_5 = 0 \end{cases} \text{ and thus}$$

$$\vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_4 \\ x_2 \\ -x_4 \\ x_4 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}.$$

We get $\text{Null space} = \text{Span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$.

<u>Useful fact</u> $\dim N(A) = $ number of free variables

$\dim N(A - \lambda I) = $ number of free variables

$= $ number of lin. indep. eigenvectors with eigenvalue λ.

Example (A 3x3 matrix) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 3 \end{bmatrix}$.

Eigenvalues: those are roots of $f(\lambda) = \det(A - \lambda I)$. We get

$$f(\lambda) = \det \begin{bmatrix} 1-\lambda & 0 & 0 \\ 1 & 1-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{bmatrix} \begin{matrix} 1-\lambda & 0 \\ 1 & 1-\lambda \\ 2 & 0 \end{matrix} = (1-\lambda)^2(3-\lambda)$$

$$\ominus \quad \ominus \quad \ominus \quad \oplus \quad \oplus \quad \oplus$$

and the eigenvalues are $\lambda = 1, 1, 3$.

Eigenvectors: those are vectors in the null space $N(A - \lambda I)$.

$\boxed{\lambda=1}$ $A - \lambda I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ ---- 1 pivot, 2 free variables

Eigenvectors satisfy $x_1 + x_3 = 0$ or $x_1 = -x_3$ so $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ x_2 \\ x_3 \end{bmatrix}$

so $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ so $N(A-I) = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

$\boxed{\lambda=3}$ $A - \lambda I = \begin{bmatrix} -2 & 0 & 0 \\ 1 & -2 & 1 \\ 2 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{bmatrix}$ ---- 2 pivots, 1 free variable

So $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2}x_3 \\ x_3 \end{bmatrix}$ and $N(A - 3I) = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$.

According to the general theorem,

$$B = [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \Rightarrow B^{-1}AB = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 3 \end{bmatrix}$$

as long as $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are lin. independent. We can check that they are. Note that we get $\underline{\underline{3}}$ eigenvectors but only $\underline{\underline{2}}$ eigenvalues.

Theorem (Distinct eigenvalues) Suppose A is $n \times n$. Eigenvectors corresponding to distinct eigenvalues are lin. independent. In particular,

$$\boxed{\begin{array}{c}A \text{ has } n\\ \text{eigenvalues}\end{array}} \Longrightarrow \boxed{\begin{array}{c}A \text{ has } n \text{ lin.}\\ \text{indep. eigenvectors}\end{array}} \Longrightarrow \boxed{\begin{array}{c}A \text{ is}\\ \text{diagonalisable}\end{array}}$$

However, we don't need n eigenvalues to ensure A is diagonalisable.

Proof. We use induction. Suppose $\vec{v}_1, \dots, \vec{v}_k$ are eigenvectors with $\lambda_1, \dots, \lambda_k$ eigenvalues distinct. When $k=1$, we get one eigenvector and $c_1 \vec{v}_1 = 0$ implies $c_1 = 0$ since $\vec{v}_1 \neq 0$ by definition. This settles $k=1$.

Assume for k and $c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1} = 0$ (*)

Multiplying by A gives $c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 + \dots + c_k\lambda_k\vec{v}_k + c_{k+1}\lambda_{k+1}\vec{v}_{k+1} = 0$.

Multiplying by λ_{k+1} gives $c_1\lambda_{k+1}\vec{v}_1 + c_2\lambda_{k+1}\vec{v}_2 + \dots + c_k\lambda_{k+1}\vec{v}_k + c_{k+1}\lambda_{k+1}\vec{v}_{k+1} = 0$.

We then get $c_1(\lambda_1 - \lambda_{k+1})\vec{v}_1 + c_2(\lambda_2 - \lambda_{k+1})\vec{v}_2 + c_k(\lambda_k - \lambda_{k+1})\vec{v}_k = 0$.

By induction $c_1(\lambda_1 - \lambda_{k+1}) = c_2(\lambda_2 - \lambda_{k+1}) = \dots = c_k(\lambda_k - \lambda_{k+1}) = 0$

$$\Rightarrow c_1 = c_2 = \dots = c_k = 0.$$

Returning to (*) gives $c_{k+1}\vec{v}_{k+1} = 0$ so $c_{k+1} = 0$ as well. ☑

Example. $A = \begin{bmatrix} 1 & 20 & 30 \\ & 2 & 4 \\ & & 3 \end{bmatrix}$ has $\lambda = 1, 2, 3 \Rightarrow A$ is diagonalisable

In particular, there exists B with $B^{-1}AB = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}$. We don't need to compute eigenvectors in this case.

Solving polynomial equations

- When A is $n \times n$, its char. polynomial $f(\lambda) = \det(A - \lambda I)$ has degree n. There are two useful facts involving such polynomials.

Theorem 1. (Rational roots) Suppose $f(x) = a_n x^n + \dots + a_2 x^2 + a_1 x + a_0$ is an integer polynomial. If $f(x)$ has a rational root x_0, then x_0 can be written as $x_0 = p/q$ with p, q relatively prime ($\gcd = 1$)
and p a divisor of a_0
and q a divisor of a_n.
In other words, one may list all possible rational roots of f.

Proof. Suppose $f(x_0) = 0$. Then $a_n \frac{p^n}{q^n} + a_{n-1} \frac{p^{n-1}}{q^{n-1}} + \dots + a_1 \frac{p}{q} + a_0 = 0$.

Clear denominators ----- $\underline{a_n p^n + a_{n-1} p^{n-1} q + \dots + a_1 p q^{n-1}} + a_0 q^n = 0$.

Since p divides the terms up until the last, p divides $a_0 q^n$ as well.
But p has no common factor with q $\Rightarrow$ p divides a_0.
Similarly, q divides $a_n p^n$ and q divides a_n. ∎

Theorem 2. (Factor theorem) If $f(x)$ is a polynomial with x_0 as a root, then $f(x)$ has $x - x_0$ as a factor, namely $f(x) = (x - x_0) g(x)$ for some g.

Proof. We divide $f(x)$ and $x - x_0$ in analogy with division of integers.

$$\begin{array}{r|l}
 & x^2 + 2x + 3 \leftarrow \text{quotient} \\
\hline
x - 2 & x^3 - x + 2 \\
 & x^3 - 2x^2 \\
\hline
 & 2x^2 - x + 2 \\
 & 2x^2 - 4x \\
\hline
 & 3x + 2 \\
 & 3x - 6 \\
\hline
 & 8 \leftarrow \text{remainder.}
\end{array}$$

$$\begin{array}{r|l}
 & Q(x) \leftarrow \text{quotient} \\
\hline
x - x_0 & f(x) \\
 & \vdots \\
 & r \leftarrow \text{remainder.}
\end{array}$$

We get $Q(x)(x - x_0) + r = f(x)$.
But x_0 is a root so $f(x_0) = 0$.
Thus $r = 0$ and there is no remainder. ∎

Example. Consider $A = \begin{bmatrix} 3 & 1 & -3 \\ 1 & 3 & -3 \\ 1 & 1 & -1 \end{bmatrix}$. Is A diagonalisable?

$\boxed{\text{Eigenvalues}}$ $\rightsquigarrow$ $f(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} 3-\lambda & 1 & -3 \\ 1 & 3-\lambda & -3 \\ 1 & 1 & -1-\lambda \end{bmatrix}$

so $f(\lambda) = (3-\lambda)^2(-1-\lambda) - 3 - 3 + 3(3-\lambda) + 3(3-\lambda) + 1 + \lambda$

$$= (9 - 6\lambda + \lambda^2)(-1-\lambda) - 6 + 18 - 6\lambda + 1 + \lambda$$

$$= -9 - 9\lambda + 6\lambda + \underbrace{6\lambda^2 - \lambda^2}\, \underline{\underline{-\lambda^3}} + 12 - 5\lambda + 1$$

$$= -\lambda^3 + 5\lambda^2 - 8\lambda + 4.$$

Possible roots $\longrightarrow$ $\lambda = \pm 1, \pm 2, \pm 4$ are only possible rational roots.

We check: $f(1) = -1 + 5 - 8 + 4 = 0$ ... so $\lambda = 1$ is a root.

$f(-1) = 1 + 5 + 8 + 4 \neq 0$ ... so $\lambda = -1$ is not a root.

$f(2) = -8 + 20 - 16 + 4 = 0$ ... so $\lambda = 2$ is a root.

Once we have a root we can factor. Since $\lambda = 1$ is a root, $\lambda - 1$ is a factor so $f(\lambda) = (\lambda - 1)\, g(x)$. We find $g(x)$ by ~~[illegible]~~ division

$$\begin{array}{r|l} & -\lambda^2 + 4\lambda - 4 \\ \hline \lambda - 1 & -\lambda^3 + 5\lambda^2 - 8\lambda + 4 \\ & \underline{-\lambda^3 + \lambda^2} \\ & \quad 4\lambda^2 - 8\lambda + 4 \\ & \quad \underline{4\lambda^2 - 4\lambda} \\ & \qquad -4\lambda + 4 \\ & \qquad \underline{-4\lambda + 4} \\ & \qquad\quad 0 \end{array}$$

This computation shows $f(\lambda) = (\lambda - 1)(-\lambda^2 + 4\lambda - 4)$

$$= (1 - \lambda)(\lambda^2 - 4\lambda + 4)$$

$$= (1 - \lambda)(\lambda - 2)^2.$$

$\boxed{\text{Eigenvectors}}$ Once we know λ, we compute $N(A-\lambda I)$. In this case

$$N(A-I) = \mathrm{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\} \quad \text{and} \quad N(A-2I) = \mathrm{Span}\left\{\begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}3\\0\\1\end{bmatrix}\right\}.$$

We get 3 eigenvectors, they are lin indep. by the theorem about distinct eigenvalues $\Rightarrow$ $B = \begin{bmatrix}1 & -1 & 3\\1 & 1 & 0\\1 & 0 & 1\end{bmatrix}$ implies $B^{-1}AB = \begin{bmatrix}1 & & \\ & 2 & \\ & & 2\end{bmatrix}$.

<u>Complex eigenvalues/eigenvectors.</u> It may happen that $\det(A-\lambda I)=0$ has complex roots. We can handle those in the same way. As an example, take $A = \begin{bmatrix}a & -b\\b & a\end{bmatrix} = \begin{bmatrix}2 & -3\\3 & 2\end{bmatrix}$, for instance.

This does not have real eigenvalues! Namely,

$$f(\lambda) = \lambda^2 - (\mathrm{tr}A)\lambda + \det A = \underbrace{\lambda^2 - 4\lambda + 4}_{} + 9 = (\lambda-2)^2 + 3^2.$$

<u>Eigenvalues</u> satisfy $(\lambda-2)^2 = -3^2 \Rightarrow \lambda - 2 = \pm 3i \Rightarrow \underline{\lambda = 2 \pm 3i}$.

<u>Eigenvectors</u>. When $\lambda = 2+3i$, we get

$$A-\lambda I = \begin{bmatrix}-3i & -3\\3 & -3i\end{bmatrix} \to \begin{bmatrix}3 & -3i\\-3i & -3\end{bmatrix} \to \begin{bmatrix}1 & -i\\0 & 0\end{bmatrix}$$

so $x - iy = 0$ and $x = iy$ and $\vec{v} = \begin{bmatrix}iy\\y\end{bmatrix} = y\begin{bmatrix}i\\1\end{bmatrix}$.

Recursive relations - A typical example

- Suppose $\begin{cases} x_n = 4x_{n-1} - 2y_{n-1} \\ y_n = x_{n-1} + y_{n-1} \end{cases}$ with $\begin{cases} x_0 = 4 \\ y_0 = 1 \end{cases}$.
- We need to determine x_n, y_n for all n. Write

$$\vec{u}_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix} \Rightarrow \vec{u}_n = \begin{bmatrix} 4x_{n-1} - 2y_{n-1} \\ x_{n-1} + y_{n-1} \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_{n-1} \\ y_{n-1} \end{bmatrix} = A \cdot \vec{u}_{n-1}$$

$$\Rightarrow \vec{u}_n = A\vec{u}_{n-1} = A^2 \vec{u}_{n-2} = A^3 \vec{u}_{n-3} = \dots = A^n \vec{u}_0.$$

We need to compute A^n when $A = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$.

- Let's diagonalise A, if possible.

<u>Eigenvalues</u> $\rightarrow$ $f(\lambda) = \lambda^2 - (\mathrm{tr}A)\lambda + \det A = \lambda^2 - 5\lambda + 6 = (\lambda-2)(\lambda-3)$

<u>Eigenvectors</u> $\rightarrow$ One has $N(A-2I) = \mathrm{Span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$ and $N(A-3I) = \mathrm{Span}\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$.

<u>General theory</u> gives $B = \left[\begin{array}{c:c} 1 & 2 \\ 1 & 1 \end{array}\right] \Rightarrow B^{-1}AB = \begin{bmatrix} 2 & \\ & 3 \end{bmatrix}$

$$\Rightarrow (B^{-1}AB)^n = \begin{bmatrix} 2^n & \\ & 3^n \end{bmatrix}$$

$$\Rightarrow B^{-1} \underline{\underline{A^n}} B = \begin{bmatrix} 2^n & \\ & 3^n \end{bmatrix}.$$

Solving for A^n gives $A^n = B \cdot \begin{bmatrix} 2^n & \\ & 3^n \end{bmatrix} \cdot B^{-1}$

and so $A^n = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2^n & \\ & 3^n \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$

$$\Rightarrow A^n = \begin{bmatrix} 2\cdot 3^n - 2^n & 2^{n+1} - 2\cdot 3^n \\ 3^n - 2^n & 2^{n+1} - 3^n \end{bmatrix}$$

and so

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \vec{u}_n = A^n \cdot \vec{u}_0 = \begin{bmatrix} 2\cdot 3^n - 2^n & 2^{n+1} - 2\cdot 3^n \\ 3^n - 2^n & 2^{n+1} - 3^n \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} 6\cdot 3^n - 2^{n+1} \\ 3^{n+1} - 2^{n+1} \end{bmatrix}.$$

Generalised eigenvectors

Suppose A is $n \times n$. If A has n lin. indep. eigenvectors, then A is diagonalisable and one can compute A^n as before ... $B = [\vec{v}_1 \dots \vec{v}_n] \Rightarrow B^{-1}AB = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

$$\Rightarrow B^{-1}A^kB = \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix}.$$

Otherwise, A will not be diagonalisable. In that case, we look for <u>generalised eigenvectors</u> .... $\vec{v} \in N(A-\lambda I)^j$ with $j \geq 2$ instead of <u>eigenvectors</u> ... $\vec{v} \in N(A-\lambda I)$.

<u>Example</u>. Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. The eigenvalues satisfy

$$f(\lambda) = \det(A-\lambda I) = \begin{bmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{bmatrix} \begin{matrix} 1-\lambda & 0 \\ 0 & 2-\lambda \\ 1 & 1 \end{matrix} = (1-\lambda)^2(2-\lambda) - (2-\lambda)$$

$$\ominus \quad \ominus \quad \ominus \quad \oplus \quad \oplus \quad \oplus$$

$$= (2-\lambda)(\cancel{1}-2\lambda+\lambda^2)$$

$$= -\lambda(2-\lambda)^2.$$

<u>Eigenvectors</u> $N(A) = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$ and $N(A-2I) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$.

This gives <u>2 linearly</u> independent eigenvectors $\Rightarrow A$ <u>not</u> diagonalisable.

We look at $N(A^2)$ and $N(A-2I)^2$. We get

$$N(A^2) = N(A) = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\} \quad \text{.... same null space}$$

$$\text{but} \quad N(A-2I)^2 = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}\right\} \quad \text{... larger null space.}$$

Using these <u>3 vectors</u> (noting that one of them is not an eigenvector) gives

$$B = \left[\begin{array}{c:c:c} -1 & 1 & 1 \\ 0 & 0 & -2 \\ 1 & 1 & 0 \end{array}\right] \Rightarrow B^{-1}AB = \begin{bmatrix} \boxed{0} & & \\ & \boxed{2} & -1 \\ & & 2 \end{bmatrix}.$$

$$\underbrace{\qquad\qquad}_{\text{eigenvectors}} \quad \underbrace{\qquad}_{\text{gen. eigenvector}}$$

This is a matrix which is almost diagonal.

Theorem (Null spaces increase & eventually stabilise) Suppose A is a square matrix with eigenvalue λ. Then the null spaces $N(A-\lambda I)$, $N(A-\lambda I)^2$, $N(A-\lambda I)^3$, ... are increasing and eventually stabilise in the sense that there is a unique integer k with

$$N(A-\lambda I)^j = N(A-\lambda I)^k \quad \text{for all } j \geq k.$$

Proof. One has

$$\begin{aligned} \vec{v} \in N(A-\lambda I)^j &\Rightarrow (A-\lambda I)^j \vec{v} = 0 \\ &\Rightarrow (A-\lambda I)(A-\lambda I)^j \vec{v} = 0 \\ &\Rightarrow v \in N(A-\lambda I)^{j+1} \text{ as well.} \end{aligned}$$

This shows that $N(A-\lambda I)^j \subseteq N(A-\lambda I)^{j+1}$ for all j.

However, the null spaces do not increase indefinitely. Namely,

$$\dim N(A-\lambda I) \leq \dim N(A-\lambda I)^2 \leq \dim N(A-\lambda I)^3 \leq \ldots \leq n$$

with $A = n \times n$. Since the dimensions are integers, we can get at most n different dimensions. We get some point at which

$$\dim N(A-\lambda I)^k = \dim N(A-\lambda I)^{k+1}$$

$$\Rightarrow \boxed{N(A-\lambda I)^k = N(A-\lambda I)^{k+1}.}$$

We claim $N(A-\lambda I)^j = N(A-\lambda I)^k \quad \forall\, j \geq k$.

This follows by induction. True when $j = k, k+1$.

- Suppose $N(A-\lambda I)^k = N(A-\lambda I)^{k+1}$ for some k.

We claim $N(A-\lambda I)^{k+1} = N(A-\lambda I)^{k+2}$ as well.

We always have $N(A-\lambda I)^{k+1} \subseteq N(A-\lambda I)^{k+2}$.

To show the opposite inclusion, we note that

$$\begin{aligned} \vec{v} \in N(A-\lambda I)^{k+2} &\Rightarrow (A-\lambda I)^{k+2} \vec{v} = 0 \\ &\Rightarrow \underbrace{(A-\lambda I)^{k+1}}\, \underbrace{(A-\lambda I)\vec{v}} = 0 \\ &\Rightarrow (A-\lambda I)\vec{v} \in N(A-\lambda I)^{k+1} = N(A-\lambda I)^k \\ &\Rightarrow (A-\lambda I)^k (A-\lambda I)\vec{v} = 0 \Rightarrow \vec{v} \in N(A-\lambda I)^{k+1}. \quad \blacksquare \end{aligned}$$

Example (Generalised eigenvectors) Let $A = \begin{bmatrix} 0 & 1 & 3 \\ -1 & 3 & 1 \\ -2 & 1 & 5 \end{bmatrix}$.

Eigenvalues satisfy $f(\lambda) = \det(A - \lambda I) = -\lambda^3 + 8\lambda^2 - 21\lambda + 18$.
Possible roots: $\pm 1, \pm 2, \pm 3, \pm 6, \pm 9, \pm 18$. We get $f(\lambda) = (2-\lambda)(\lambda-3)^2$.

$\boxed{\lambda = 2}$ gives $N(A-2I) = \operatorname{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\right\}$ and $N(A-2I)^2 = N(A-2I)$.
This implies $N(A-2I)^j = N(A-2I)$ $\forall j \geq 1$.

$\boxed{\lambda = 3}$ gives $N(A-3I) = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$ and $N(A-3I)^2 = \operatorname{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$
and $N(A-3I)^3 = N(A-3I)^2$. Thus $N(A-3I)^j = N(A-3I)^2$ $\forall j \geq 2$.

This gives 1 eigen. with $\lambda = 2$, 1 eigen. with $\lambda = 3$ $\Rightarrow$ 2 lin. indep. eigenvectors $\Rightarrow$ A is not diagonalisable. However, there are 3 lin. indep. generalised eigenvectors.

Our goal: to show that we always get n lin. indep. gen. eigenvectors. This is hard to prove. We'll need several concepts.

$\boxed{\text{Column space}}$

If A is $m \times n$, then
$$C(A) = \text{span of columns of } A = \operatorname{Span}\{A\vec{e}_1, A\vec{e}_2, \ldots, A\vec{e}_n\}.$$

Note that $\vec{y} \in C(A)$ if and only if $\vec{y} = x_1 A\vec{e}_1 + \ldots + x_n A\vec{e}_n$,
so $\vec{y} \in C(A) \Leftrightarrow \vec{y} = \sum x_i A\vec{e}_i = A \underbrace{\sum x_i \vec{e}_i} \Leftrightarrow \vec{y} = A \cdot \vec{x}$.

Finding a basis We can relate $C(A)$ with $C(R)$, the column space of the RREF, as follows. Suppose $R = E_k \cdots E_2 E_1 A$. Then
linear combinations of columns of A ... $\sum x_i A\vec{e}_i$
are related to linear comb. of columns of R since
$$\underbrace{\sum x_i R\vec{e}_i} = \sum x_i (E_k \cdots E_2 E_1 A)\vec{e}_i = (E_k \cdots E_1) \underbrace{\sum x_i A\vec{e}_i}.$$

If some columns of R are linearly dependent, then

$$x_1 R\vec{e}_{i_1} + x_2 R\vec{e}_{i_2} + \dots + x_k R\vec{e}_{i_k} = 0 \quad \text{with } x_j \text{ not all zero}$$

and so

$$x_1 A\vec{e}_{i_1} + x_2 A\vec{e}_{i_2} + \dots + x_k A\vec{e}_{i_k} = 0 \quad \text{(and vice versa).}$$

Theorem (Column space)

① Lin. indep. columns of R correspond to lin. indep. columns of A.

② $\dim N(A)$ = number of free variables

$\dim C(A)$ = number of pivots

So $\dim N(A) + \dim C(A)$ = number of variables / columns $= n$.

Proof. ① follows by above. For ②, row reduction gives some pivot variables and some free variables. We know

$$\dim N(A) = \text{number of free variables.}$$

Also,

$$\dim C(A) = \dim \operatorname{Span}\{A\vec{e}_1, \dots, A\vec{e}_n\}$$
$$= \dim \operatorname{Span}\{R\vec{e}_1, \dots, R\vec{e}_n\} \quad \text{by ①.}$$

However, R = the reduced row echelon form and the linearly independent columns are those that contain pivots (the other ones are linear combinations of those). For instance,

$$R = \begin{bmatrix} \boxed{1} & 2 & \boxed{0} & 2 & \boxed{0} & 0 & 5 \\ \boxed{0} & 0 & \boxed{1} & 3 & \boxed{0} & 0 & 2 \\ \boxed{0} & 0 & \boxed{0} & 0 & \boxed{1} & 0 & 1 \end{bmatrix}.$$

The columns with pivots are lin. indep. because they are the standard basis vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. The remaining columns do not contain pivots, so every nonzero entry in those columns has a pivot to its left. Thus, the column itself is a lin. comb. of the pivot columns. ▨

<u>Example</u>. Let $A = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 2 & 0 \end{bmatrix}$.

In this case $R = \begin{bmatrix} 1 & 0 & \boxed{1} & 0 & \boxed{-1} \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$.

Thus, $R\vec{e}_1$, $R\vec{e}_2$ and $R\vec{e}_4$ --- columns with pivots (lin. indep. columns)

so $A\vec{e}_1$, $A\vec{e}_2$ and $A\vec{e}_4$ are lin. indep. as well.

Also $R\vec{e}_3 = R\vec{e}_1 - R\vec{e}_2$ and so $A\vec{e}_3 = A\vec{e}_1 - A\vec{e}_2$.

<u>Theorem</u> (Coordinate vectors) Let A be $n \times n$ and $B = [\vec{v}_1\ \vec{v}_2 \dots \vec{v}_n]$. Then the $k^{\underline{th}}$ column of $B^{-1}AB$ lists the coefficients one needs to express $A\vec{v}_k$ in terms of the basis $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$.

<u>Proof</u>. $B^{-1}AB\vec{e}_k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \iff B^{-1}AB\vec{e}_k = \sum_{i=1}^{n} x_i \vec{e}_i$

$$\iff A\underline{B\vec{e}_k} = \sum x_i \underline{B\vec{e}_i} \iff A\vec{v}_k = \sum x_i \vec{v}_i. \quad \blacksquare$$

<u>Remark</u>. To ensure $B^{-1}AB$ is as simple as possible, with as many zeros as possible, we need $A\vec{v}_k$ to be expressible in terms of a small number of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$. We've seen eigenvectors ... $A\vec{v}_k = \lambda\vec{v}_k$.

<u>Example</u> Suppose A is 3×3 and $B = [\vec{v}_1\ \vec{v}_2\ \vec{v}_3]$. Suppose also

$$\left\{ \begin{aligned} Av_1 &= \lambda v_1 \\ Av_2 &= a_1 v_2 + a_2 v_3 \\ Av_3 &= b_1 v_2 + b_2 v_3 \end{aligned} \right\}.$$

Then $B^{-1}AB = \left[\begin{array}{c|c:c} \lambda & 0 & 0 \\ \hline 0 & a_1 & b_1 \\ 0 & a_2 & b_2 \end{array}\right]$ by the theorem.

We get a 1×1 block $[\lambda]$ corresponding to $\vec{v}_1$

and a 2×2 block $\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$ corresponding to $\vec{v}_2, \vec{v}_3$.

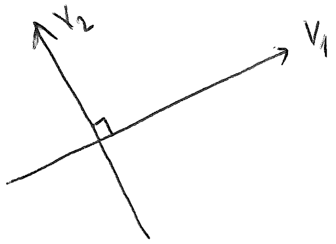
Invariant subspaces

Definition We say $\mathcal{U}$ is an A-invariant subspace for some matrix A, if $A\vec{u} \in \mathcal{U}$ for every vector $\vec{u} \in \mathcal{U}$. Thus, left multiplication by A leaves $\mathcal{U}$ invariant.

One-dimensional invariant subspaces $\mathcal{U} = \operatorname{Span}\{\vec{v}\}$

Invariance means $A\vec{v} = \lambda\vec{v}$ so we get eigenvectors $\vec{v}$.

Reflections in $\mathbb{R}^2$



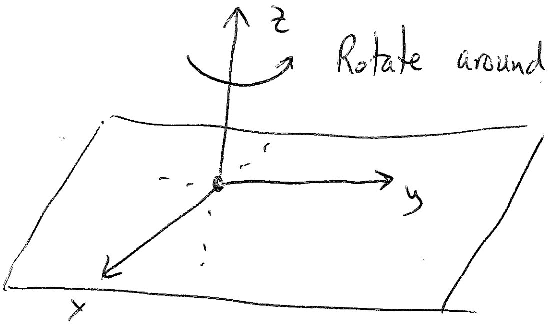
Suppose $T(\vec{x}) = A\vec{x}$ is reflection along a line in $\mathbb{R}^2$.

Then $A\vec{v}_1 = \vec{v}_1$ ---- 1-dimensional inv. subspace

and $A\vec{v}_2 = -\vec{v}_2$ ---- 1-dimensional inv. subspace.

Thus $B = [\vec{v}_1\ \vec{v}_2] \Rightarrow B^{-1}AB = \left[\begin{array}{c|c} 1 & \\ \hline & -1 \end{array}\right]$.

Rotations in $\mathbb{R}^3$



Suppose $T(\vec{x}) = A\vec{x}$ is rotation around some axis/line in $\mathbb{R}^3$, say the z-axis.

Then $A\vec{e}_3 = \vec{e}_3$ ---- 1-dimens. inv. subsp.

and $\operatorname{Span}\{e_1, e_2\}$ --- 2-dim. inv. subspace.

In this case $B^{-1}AB = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & * & * \\ 0 & * & * \end{array}\right]$.

Theorem (Null spaces and column spaces) Let A be $n \times n$ and suppose λ is an eigenvalue, possibly complex. Then

$N(A - \lambda I)^j$ is an A-invariant subspace of $\mathbb{C}^n$,

namely $\vec{v} \in N(A - \lambda I)^j \Rightarrow A\vec{v} \in N(A - \lambda I)^j$ as well.

Moreover, $C(A - \lambda I)^j$ is A-invariant as well.

Proof. Let's worry about column spaces. To say $\vec{y} \in C(B)$ is to say $\vec{y} = \sum x_i B\vec{e}_i$ and that says $\vec{y} = B\vec{x}$. In our case, $\vec{y} \in C(A-\lambda I)^j$

$$\Rightarrow \vec{y} = (A-\lambda I)^j \vec{x}$$

$$\Rightarrow A\vec{y} = A(A-\lambda I)^j \vec{x}$$

$$\Rightarrow A\vec{y} = (\underbrace{A-\lambda I}+\underbrace{\lambda I})(A-\lambda I)^j \vec{x}$$

$$\Rightarrow A\vec{y} = (A-\lambda I)^{j+1}\vec{x} + \lambda(A-\lambda I)^j \vec{x}$$

$$= (A-\lambda I)^j \vec{w}. \quad \blacksquare$$

Example (Block diagonalisation) Let $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 0 & 2 \\ 0 & -1 & 2 \end{bmatrix}$.

The eigenvalues are ... $\lambda = 1, 1, 2$.

The eigenvectors are ... $N(A-I) = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ and $N(A-2I) = \text{Span}\left\{\begin{bmatrix}2\\0\\1\end{bmatrix}\right\}$.

Thus, A is not diagonalisable. The generalised eigenvectors are

$N(A-I)^2 = \text{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$ and $N(A-2I)^2 = N(A-2I)$.

We get 2 gen. eigenv. with $\lambda = 1$ and 1 eigenv. with $\lambda = 2$.

Thus $N(A-I)^2$ = 2-dim. inv. subspace and $N(A-2I)^2$ = 1-dim. inv.

We take $B = \left[v_1 \; v_2 \,\middle|\, v_3 \right] = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ and then

compute $B^{-1}AB = \left[\begin{array}{cc|c} 1 & -1 & \\ 0 & 1 & \\ \hline \multicolumn{2}{c|}{0} & 2 \end{array}\right]$ ... a block diagonal matrix.