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
  \[
  \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 4x_{n-1} - 2y_{n-1} \\ x_{n-1} + y_{n-1} \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = A \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix},
  \]
  \[
  \Rightarrow \begin{pmatrix} x_n \\ y_n \end{pmatrix} = A^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = A^2 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = A^3 \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \ldots = A^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.
  \]
  We need to compute \( A^n \) when \( A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \).

• Let's diagonalise \( A \), if possible.
  Eigenvalues \( \rightarrow f(\lambda) = \lambda^2 - (tr A) \lambda + \det A = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3) \)
  Eigenvectors \( \rightarrow \) one has \( N(A-2I) = \text{Span} \{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \} \) and \( N(A-3I) = \text{Span} \{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \} \).
  General theory gives \( B = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \) \( \Rightarrow \) \( B^{-1} A B = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \)
  \[
  \Rightarrow (B^{-1} A B)^n = \begin{pmatrix} 2^n & 3^n \\ 1 & 1 \end{pmatrix}.
  \]

  Solving for \( A^n \) gives \( A^n = B \begin{pmatrix} 2^n & 3^n \\ 1 & 1 \end{pmatrix} B^{-1} \).

and so \( A^n = \begin{pmatrix} 4 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^n & 3^n \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \)

\[
\Rightarrow A^n = \begin{pmatrix} 2 \cdot 3^n - 2^n & 2^{n+1} - 2 \cdot 3^n \\ 3^n - 2^n & 2^{n+1} - 3^n \end{pmatrix}.
\]

and so

\[
\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = A^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3^n - 2^n & 2^{n+1} - 2 \cdot 3^n \\ 3^n - 2^n & 2^{n+1} - 3^n \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} \]

\[
\Rightarrow \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 6 \cdot 3^n - 2^{n+1} \\ 3^{n+1} - 2^{n+1} \end{pmatrix}.
\]
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 ... $B = [v_1 \ldots v_n] \Rightarrow B^{-1}AB = [\lambda_1 \ldots \lambda_n]$

\[ B^{-1}A^kB = [\lambda_1^k \ldots \lambda_n^k]. \]

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

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

\[ f(\lambda) = \det (A-\lambda I) = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2(2-\lambda) - (2-\lambda) = (2-\lambda)(\lambda^2 - 6\lambda + 5) \]

\[ = -\lambda(2-\lambda)^2. \]

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

This gives 2 linearly independent eigenvectors $= A$ not diagonalisable.

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

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

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

Using these 3 vectors (noting that one of them is not an eigenvector)

\[ B = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & -2 \\ 1 & 1 & 0 \end{bmatrix} \Rightarrow B^{-1}AB = \begin{bmatrix} 0 & 2 & -1 \\ \end{bmatrix} \]

\begin{bmatrix} 0 & 2 & -1 \\ \end{bmatrix}

\begin{bmatrix} \text{eigenvectors} \quad \text{gen.eigenvector} \\ \end{bmatrix}

This is a matrix which is almost diagonal.
Theorem: (Null spaces increase & eventually stabilise) Suppose $A$ is a square matrix with eigenvalue $\lambda$. 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 an integer $k$ with

$$N(A - \lambda I)^j = N(A - \lambda I)^k$$

for all $j \geq k$.

Proof. One has $\forall \mathbf{v} \in N(A - \lambda I)^3 \Rightarrow (A - \lambda I)^3 \mathbf{v} = 0$

$$\Rightarrow (A - \lambda I)(A - \lambda I)^3 \mathbf{v} = 0$$

$$\Rightarrow \mathbf{v} \in N(A - \lambda I)^4$$

as well.

This shows that $N(A - \lambda I)^3 \subseteq N(A - \lambda I)^4$ 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

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

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

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

$$\mathbf{v} \in N(A - \lambda I)^{k+2} \Rightarrow (A - \lambda I)^{k+2} \mathbf{v} = 0$$

$$\Rightarrow (A - \lambda I)^{k+1} (A - \lambda I) \mathbf{v} = 0$$

$$\Rightarrow (A - \lambda I) \mathbf{v} \in N(A - \lambda I)^{k+1} = N(A - \lambda I)^k$$

$$\Rightarrow (A - \lambda I)^k (A - \lambda I) \mathbf{v} = 0 \Rightarrow \mathbf{v} \in N(A - \lambda I)^{k+1}$$
Example (Generalised eigenvectors) Let \( A = \begin{bmatrix} -1 & 1 & 5 \\ 2 & 1 & 5 \end{bmatrix} \).

Eigenvalues satisfy \( f(\lambda) = \det(A - \lambda I) = -\lambda^3 + 8\lambda^2 - 24\lambda + 18 \).
Possible roots: \( \pm 1, \pm 2, \pm 3, \pm 6, \pm 9, \pm 18 \). We get \( f(\lambda) = (2-\lambda)(\lambda-3)^2 \).

\( \lambda = 2 \) gives \( N(A - 2I) = \text{Span} \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\} \) and \( N(A - 2I)^2 = N(A - 2I) \).
This implies \( N(A - 2I)^3 = N(A - 2I) \) \( \forall j \geq 1 \).

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

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.

**Column space**

If \( A \) is \( m \times n \), then \( C(A) = \text{Span of columns of } A \)
\[ = \text{Span} \left\{ A\vec{e}_1, A\vec{e}_2, \ldots, A\vec{e}_n \right\} \]

Note that \( \vec{y} \in C(A) \) if and only if \( \vec{y} = x_1 A\vec{e}_1 + \cdots + x_n A\vec{e}_n \),
so \( \vec{y} \in C(A) \iff \vec{y} = \sum x_i A\vec{e}_i = A \sum x_i \vec{e}_i \iff \vec{y} = A \times \).

Finding a basis: We can relate \( C(A) \) with \( C(R) \), the column space of the RREF, as follows. Suppose \( R = \vec{e}_1 \ldots \vec{e}_k \vec{A} \). Then linear combinations of columns of \( A \)
\[ \sum x_i \vec{A} \vec{e}_i \]
are related to linear comb. of columns of \( R \) since
\[ \sum x_i \vec{e}_i \vec{R} \vec{e}_i = \sum x_i \left( \vec{e}_1 \ldots \vec{e}_k \vec{A} \right) \vec{e}_i = \left( \vec{e}_1 \ldots \vec{e}_k \right) \sum x_i \vec{A} \vec{e}_i. \]
If some columns of $A$ are linearly dependent, then
\[ x_1 \mathbf{R} \mathbf{e}_1 + x_2 \mathbf{R} \mathbf{e}_2 + \ldots + x_k \mathbf{R} \mathbf{e}_k = 0 \quad \text{with } x_j \text{ not all zero} \]
and so
\[ x_1 \mathbf{A} \mathbf{e}_1 + x_2 \mathbf{A} \mathbf{e}_2 + \ldots + x_k \mathbf{A} \mathbf{e}_k = 0 \quad (\text{and vice versa}). \]

**Theorem (Column space)**

1. Lin. indep. columns of $A$ correspond to lin. indep. columns of $A$.
2. $\dim N(A) = \text{number of free variables}$
   \[ \dim \mathcal{C}(A) = \text{number of pivots} \]
   \[ \quad \Rightarrow \quad \dim N(A) + \dim \mathcal{C}(A) = \text{number of variables/columns} = n. \]

**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,
\[ \dim \mathcal{C}(A) = \dim \text{Span } \{ \mathbf{A} \mathbf{e}_1, \ldots, \mathbf{A} \mathbf{e}_n \} \]
\[ = \dim \text{Span } \{ \mathbf{R} \mathbf{e}_1, \ldots, \mathbf{R} \mathbf{e}_n \} \quad \text{by (1)}. \]

However, $R$ is 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,
\[
R = \begin{bmatrix}
1 & 2 & 0 & 2 & 0 & 0 & 5 \\
0 & 0 & 1 & 3 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]

The columns with pivots are lin. indep. because they are the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. 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.
Example: Let \( A = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 & 0 \end{bmatrix} \).

In this case \( R = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \).

Thus, \( \mathbf{R} e_1, \mathbf{R} e_2 \) and \( \mathbf{R} e_4 \) are columns with pivots (lin. indep. columns)

so \( A e_1, A e_2 \) and \( A e_4 \) are lin. indep. as well.

Also \( \mathbf{R} e_3 = \mathbf{R} e_1 - \mathbf{R} e_2 \) and so \( A e_3 = A e_1 - A e_2 \).

**Theorem (Coordinate vectors)** Let \( A \) be \( n \times n \) and \( B = [\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_n] \).

Then the \( k \text{th} \) column of \( B^{-1} A B \) lists the coefficients one needs to express \( A \mathbf{v}_k \) in terms of the basis \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \).

**Proof.** \( B^{-1} A \mathbf{e}_k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \iff B^{-1} A \mathbf{e}_k = \Sigma x_i \mathbf{e}_i \)

\( \iff A \mathbf{b}_k = \Sigma x_i \mathbf{b}_i \iff A \mathbf{v}_k = \Sigma x_i \mathbf{v}_i \).

**Remark.** To ensure \( B^{-1} A B \) is as simple as possible, with as many zeros as possible, we need \( A \mathbf{v}_k \) to be expressible in terms of a small number of the vectors \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \). We've seen eigenvectors \( A \mathbf{v}_k = \lambda \mathbf{v}_k \).

**Example** Suppose \( A \) is \( 3 \times 3 \) and \( B = [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3] \). Suppose also

\[
\begin{cases}
A \mathbf{v}_1 = 2 \mathbf{v}_1 \\
A \mathbf{v}_2 = a_1 \mathbf{v}_2 + a_2 \mathbf{v}_3 \\
A \mathbf{v}_3 = b_1 \mathbf{v}_2 + b_2 \mathbf{v}_3
\end{cases}
\]

Then \( B^{-1} A B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & a_1 & b_1 \\ 0 & a_2 & b_2 \end{bmatrix} \) by the theorem.

We get a \( 1 \times 1 \) block \( 2 \) corresponding to \( \mathbf{v}_1 \)

and a \( 2 \times 2 \) block \( \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \) corresponding to \( \mathbf{v}_2, \mathbf{v}_3 \).
**Invariant subspaces**

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

**One-dimensional invariant subspaces**

\[ U = \text{Span} \{ \vec{v} \} \]

Invariance means \( A \vec{v} = \lambda \vec{v} \), so we get eigenvectors \( \vec{v} \).

Suppose \( T(x) = A \vec{x} \) 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 \( B = [\vec{v}_1 \, \vec{v}_2] \Rightarrow B^{-1}AB = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \).

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

Then \( A \vec{e}_3 = \vec{e}_3 \) --- 1-dim. inv. subspace.

and \( \text{Span} \{ \vec{e}_1, \vec{e}_2 \} \) --- 2-dim. inv. subspace.

In this case \( B^{-1}AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \).

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

\[ N(A - \lambda I)^3 \] is an \( A \)-invariant subspace of \( \mathbb{C}^n \), namely \( \vec{v} \in N(A - \lambda I)^3 \Rightarrow A \vec{v} \in N(A - \lambda I)^3 \) as well.

Moreover, \( (CA - \lambda I)^3 \) is \( A \)-invariant as well.
Proof. Let’s worry about column spaces. To say \( y \in C(B) \) is to say \( y = \sum x_i \vec{e}_i \) and that says \( y = B \vec{x} \). In our case, \( y \in C(A-\lambda I) \Rightarrow y = (A-\lambda I)^3 \vec{x} \Rightarrow A_y = A (A-\lambda I)^2 \vec{x} \Rightarrow A_y = (A-\lambda I)^3 \vec{x} + \lambda (A-\lambda I)^2 \vec{x} = (A-\lambda I)^3 \vec{w} \). \( \square \)

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 \( N(A-I) = \text{Span} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \) and \( N(A-2I) = \text{Span} \begin{bmatrix} 2 \\ 0 \end{bmatrix} \). Thus, \( A \) is not diagonalisable. The generalised eigenvectors are

\[
N(A-I) = \text{Span} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad N(A-2I) = N(A-2I).
\]

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 \( B = \begin{bmatrix} V_1 & V_2 & V_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \) and then compute \( B^{-1}AB = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix} \) — a block diagonal matrix.