5.2 Review of eigenvectors and eigenvalues

- λ is an eigenvalue of $\mathbf{A}$ if determinant $|\mathbf{A} - \lambda\mathbf{I}| = 0$
- This determinant is a polynomial in λ of degree n: so it has n roots $\lambda_1, \lambda_2, \ldots, \lambda_n$
- Every symmetric matrix $\mathbf{A}$ has a full set (basis) of n orthogonal unit eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$
- No algebraic formula for the polynomial roots for $n > 4$
  - Thus, the eigenvalue problem needs own special algorithms
  - Solving the eigenvalue problem is harder than solving $\mathbf{Ax} = \mathbf{b}$
- Determinant $|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i = \lambda_1\lambda_2 \cdots \lambda_n$ (the product of eigenvalues)
- The *trace* of a matrix is the sum of the diagonal elements. That is, $\text{trace}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn}$.
- It turns out that $\text{trace}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 + \ldots + \lambda_n$ (the sum of eigenvalues)
- $\mathbf{A}^k = \underbrace{\mathbf{A} \cdots \mathbf{A}}_{k \text{ times}}$ has the same eigenvectors as $\mathbf{A}$: e.g. for $\mathbf{A}^2$

$$\mathbf{Ae} = \lambda\mathbf{e} \;\Rightarrow\; \mathbf{AAe} = \lambda\mathbf{Ae} = \lambda^2\mathbf{e}$$

- Eigenvalues of $\mathbf{A}^k$ are $\lambda_1^k, \ldots, \lambda_n^k$
- Eigenvalues of $\mathbf{A}^{-1}$ are $\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$

Example: Find the eigenvalues and eigenvectors of $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

Solution: First, find the eigenvalues of $\mathbf{A}$ by solving

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0.$$

So the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 3$

The eigenvector associated with $\lambda_1 = 1$ is $\mathbf{e}_1$ and satisfies $\mathbf{Ae}_1 = \lambda_1\mathbf{e}_1$. Putting $\mathbf{e}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ we need to solve

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}.$$

The second row gives $-x_1 + 2y_1 = y_1$, so $y_1 = x_1$. So fix $x_1 = 1$ and $e_1 = c \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for any $c \neq 0$. If we choose c so that e_1 is normalised, $\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. A similar argument shows $\mathbf{e}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Before leaving this example, it is worth looking at some of the properties of the eigenvalues of $\mathbf{A}$:

- Determinant $\det \mathbf{A} \equiv |\mathbf{A}| = 4 - 1 = 3 \Longleftrightarrow \lambda_1 \cdot \lambda_2 \equiv 1 \cdot 3 = 3$
- $\mathrm{trace}(\mathbf{A}) = 2 + 2 = 4 \Longleftrightarrow \lambda_1 + \lambda_2 \equiv 1 + 3 = 4$
- Inverse matrix $\mathbf{A}^{-1} = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$: eigenvalues $\lambda_1 = \frac{1}{3}$ and $\lambda_2 = 1$
- Matrix $\mathbf{A}^2 = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$: eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 9$
- Matrix $\mathbf{A}^3 = \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix}$: eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 27$

$\square$

7.1 LU decomposition via Gaussian elimination

7.1.1 Gaussian elimination to solve systems linear equations (review)

You should be familiar with the process of *Gaussian elimination* (or *row reduction*) in which the equation $\mathbf{Ax} = \mathbf{b}$ (where A is arbitrary) in transformed into the equivalent equation $\mathbf{C}x = \mathbf{d}$ where $\mathbf{C}$ is triangular, making the equation easy to solve. We review the process here.

It is easy to show that *multiplying* both sides of $\mathbf{Ax} = \mathbf{b}$ from the left by any nonsingular matrix $\mathbf{M}$ does not affect the solution. That is $\mathbf{MAx} = \mathbf{Mb}$ has the same solution as $\mathbf{Ax} = \mathbf{b}$, since

$$\mathbf{MAx} = \mathbf{Mb} \Rightarrow \mathbf{x} = (\mathbf{MA})^{-1}\mathbf{Mb} = \mathbf{A}^{-1}\mathbf{M}^{-1}\mathbf{Mb} = \mathbf{A}^{-1}\mathbf{b}.$$

We know from the above result that we can multiple both sides by a series of *elementary matrices* which perform various *row operations* on $\mathbf{A}$: the three types of operation are row swapping, row multiplication and adding some multiple of one row to another row. Repeated application of these three operations (that is, repeated multiplication by elementary matrices) to both sides of the equation transforms it to $\mathbf{C}x = \mathbf{d}$ where $\mathbf{C} = \mathbf{M}_1 \ldots \mathbf{M}_k\mathbf{A}$ is in upper triangular form (so the only non-zero elements of $\mathbf{C}$ are on or above the diagonal) and $\mathbf{d} = \mathbf{M}_1 \ldots \mathbf{M}_k\mathbf{b}$.

Example: Use Gaussian elimination to solve the system of equations $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 & 2 \\ 6 & 6 & 3 & 5 \\ 3 & 0 & 3 & 5 \\ 9 & 2 & 7 & 8 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 4 \\ 5 \\ 5 \\ 10 \end{bmatrix}$$

Solution:

$$\overbrace{\begin{bmatrix} 3 & 2 & 1 & 2 \\ 6 & 6 & 3 & 5 \\ 3 & 0 & 3 & 5 \\ 9 & 2 & 7 & 8 \end{bmatrix}}^{\mathbf{A}} \overbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}^{\mathbf{x}} = \overbrace{\begin{bmatrix} 4 \\ 5 \\ 5 \\ 10 \end{bmatrix}}^{\mathbf{b}}$$

$$\Rightarrow \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 & 2 \\ 6 & 6 & 3 & 5 \\ 3 & 0 & 3 & 5 \\ 9 & 2 & 7 & 8 \end{bmatrix}}_{\text{Eliminating the first column: } \mathbf{M}_1\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 5 \\ 10 \end{bmatrix}}_{\mathbf{M}_1\mathbf{b}}$$

$$\Rightarrow \begin{bmatrix} 3 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & -2 & 2 & 3 \\ 0 & -4 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ 1 \\ -2 \end{bmatrix}$$

$$\Rightarrow \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & -2 & 2 & 3 \\ 0 & -4 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{M}_2\mathbf{M}_1\mathbf{A}\mathbf{x}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \\ 1 \\ -2 \end{bmatrix}}_{\mathbf{M}_2\mathbf{M}_1\mathbf{b}}$$

$$\Rightarrow \begin{bmatrix} 3 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 6 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ -2 \\ -8 \end{bmatrix}$$

$$\Rightarrow \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 6 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{M}_3\mathbf{M}_2\mathbf{M}_1\mathbf{A}\mathbf{x}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \\ -2 \\ -8 \end{bmatrix}}_{\mathbf{M}_3\mathbf{M}_2\mathbf{M}_1\mathbf{b}}$$

$$\Rightarrow \begin{bmatrix} 3 & 2 & 1 & 2 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \\ -2 \\ -4 \end{bmatrix}$$

It is easy to see that the solution to this row reduced matrix equation is

$$\begin{aligned} x_4 &= \tfrac{-4}{-4} &&= 1 \\ x_3 &= \tfrac{1}{3}(-2 - 4 \cdot 1) &&= -2 \\ x_2 &= \tfrac{1}{2}(-3 - 1 \cdot (-2) - 1 \cdot 1) &&= -1 \\ x_1 &= \tfrac{1}{3}(4 - 2 \cdot (-1) - 1 \cdot (-2) - 2 \cdot 1) &&= 2 \end{aligned}$$

□

7.1.2 Gaussian elimination as LU decomposition

It turns out that Gaussian elimination can be viewed a LU decomposition in which a matrix $\mathbf{A}$ is written as the product of a lower triangular matrix $\mathbf{L}$ and an upper triangular matrix $\mathbf{U}$ so that $\mathbf{A} = \mathbf{LU}$. Recall that the Gaussian elimination process starts with an arbitrary square matrix $\mathbf{A}$, multiplies it by a series of elementary vectors, $\mathbf{M}_1 \ldots \mathbf{M}_k$ to get $\mathbf{U} = \mathbf{M}_1 \ldots \mathbf{M}_k\mathbf{A}$ where $\mathbf{U}$ is upper triangular (we called it $\mathbf{C}$ in the earlier discussion).

Now (check that you understand the following statements), each of the elementary matrices is lower triangular (so long as there are no row swapping operations) and the inverse of lower triangular matrices are lower triangular too, so each of $\mathbf{M}_i^{-1}$, for $i = 1, \ldots, k$ is lower triangular. Finally, the product of lower triangular matrices is also lower triangular so

$$\mathbf{U} = \mathbf{M}_1 \ldots \mathbf{M}_k \mathbf{A} \implies \mathbf{LU} = \mathbf{A},$$

where $\mathbf{L} = (\mathbf{M}_1 \ldots \mathbf{M}_k)^{-1} = \mathbf{M}_k^{-1} \ldots \mathbf{M}_1^{-1}$. If row permutations (row swaps) are needed in the Gaussian elimination process, we can't find an LU decomposition for $\mathbf{A}$ but can find an LU decomposition for the permuted matrix $\mathbf{PA}$, where $\mathbf{P}$ describes the necessary permutations. Obviously, the permuted system has the same solution as the unpermuted system.

The computational complexity of solving a system of n equations in n unknown using Gaussian elimination is $O(n^3)$. It is typically a stable algorithm, though potential for instability arises when a leading non-zero entry is very small (as we divide through by this entry). Reordering of the rows before the start of the row reduction process so that the largest leading non-zero elements are selected first can avoid this cause of instability. This technique is known as pivoting.

We won't look further at LU decompositions but will spend considerable time looking first at singular value decomposition and its uses and, later, when we consider the Least Squares framework, QR decompositions and its applications.