Last time: Eigenvalues & Eigenvectors

Def: Let $A$ be an $n \times n$ matrix. A scalar $\lambda$ and a non-zero vector $v$ is called an eigenvalue/eigenvector pair if $Av = \lambda v$.

Def: The characteristic polynomial of an $n \times n$ matrix $A$ is $p(\lambda) = \det(A - \lambda I_n)$.

Fact: The roots of $p(\lambda)$ are the eigenvalues of $A$.

Def: Let $A$ be an $n \times n$ matrix. Then we call $W(A - \lambda I_n) = \mathbb{E}_\lambda$ the eigenspace corresponding to $\lambda$.

Fact: Eigenvectors corresponding to an eigenvalue $\lambda$ of an $n \times n$ matrix $A$ are given by the basis vectors of $\mathbb{E}_\lambda$.

Remark: Eigenvectors corresponding to different eigenvalues are linearly independent.

Fact: Let $A$ be an $n \times n$ matrix. If all $n$ eigenvalues are distinct, then the corresponding eigenvectors form a basis for $\mathbb{R}^n$ (resp. $\mathbb{C}^n$).

Def: i) The algebraic multiplicity of an eigenvalue $\lambda$ of a matrix $A$ is the number of times that $\lambda$ appears as a root of the characteristic polynomial of $A$.

ii) The geometric multiplicity of an eigenvalue $\lambda$ of a matrix $A$ is the dimension of the eigenspace $\mathbb{E}_\lambda$.

Fact: There exists an eigenbasis corresponding to the matrix $A$ if and only if the algebraic multiplicity is equal to the geometric multiplicity for each eigenvalue $\lambda$ of $A$.

Fact: If $A \in \mathbb{R}^{n \times n}$ and $\lambda$ is an eigenvalue of $A$ with eigenvector $v$, then $\lambda$ is also an eigenvalue of $A$ with eigenvector $v$.
Last time continued: Diagonalization

Setting: Let \( A \) be \( n \times n \) with eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \) and an eigenbasis \( \{v_1, v_2, \ldots, v_n\} \) such that \( Av_j = \lambda_j v_j \).

Def.: A square matrix \( A \) is diagonalizable if an invertible matrix \( S \) and a diagonal matrix \( D \) exist such that \( \tilde{A} = SDS^{-1} \).

Fact. Let \( S = [v_1 \ v_2 \ldots \ v_n] \) and \( D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \). Then \( S \) is invertible since \( \{v_1, v_2, \ldots, v_n\} \) is an eigenbasis. Moreover \( A \) is diagonalizable since \( A = SDS^{-1} \).

Fact. i) The determinant of a diagonalizable matrix is equal to the product of its eigenvalues.
   ii) The trace of a diagonalizable matrix is equal to the sum of its eigenvalues.
   iii) For \( \lambda \in \mathbb{N} \) and a diagonalizable matrix \( A \) which is \( n \times n \) we have that \( S^\lambda S^{-1} = A^\lambda \) where \( S \) and \( D \) are defined as above and \( \lambda_1, \lambda_2, \ldots, \lambda_n \) are the eigenvalues of \( A \).
   iv) If all eigenvalues \( \lambda_j \) in iii) are non-zero then this fact also holds for the negative whole numbers.

Hermitian matrices

Def.: A square matrix \( A \) is called Hermitian if \( A^* = A \).

Remark: i) All diagonal entries of Hermitian matrices are real.
   ii) A Hermitian matrix that has only real entries is called symmetric.

Fact. i) \( A \) is Hermitian iff \( \langle v, Aw \rangle = \langle Av, w \rangle \) for all \( v, w \).
   ii) Eigenvalues of Hermitian matrices are real.
   iii) Eigenvectors of Hermitian matrices corresponding to distinct eigenvalues are orthogonal.
   iv) Any Hermitian matrix is unitarily diagonalizable, i.e., \( A = UDU^* \).
   v) It is easy to define powers and other functions of Hermitian matrices.
Power method for computing eigenvalues

Goal: Find a single eigenvalue-eigenvector pair

Remark: It will turn out that this method gives us the eigenvalue with the biggest absolute value.

Suppose $A$ is $n \times n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding eigenvectors $v_1, \ldots, v_n$ such that

(i) $|\lambda_1| > |\lambda_2| > |\lambda_3| > \ldots > |\lambda_n|$

(ii) $\{v_1, \ldots, v_n\}$ is an eigenbasis

(iii) $\|v_j\|_2 = 1$, $j = 1, \ldots, n$.

For simplicity we also assume that $\lambda_j \in \mathbb{R}$. This is the case if $A$ is real-symmetric.

Now pick an arbitrary $x_0 \in \mathbb{R}^n$ with $x_0 \neq 0$. Then there exist $c_1, \ldots, c_n$ such that

$$x_0 = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$

Moreover, from this representation as a linear combination we can deduce

$$A x_0 = c_1 A v_1 + c_2 A v_2 + \ldots + c_n A v_n = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_n \lambda_n v_n$$

$$A^2 x_0 = c_1 \lambda_1^2 v_1 + c_2 \lambda_2^2 v_2 + \ldots + c_n \lambda_n^2 v_n$$

\[ \vdots \]

$$A^k x_0 = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \ldots + c_n \lambda_n^k v_n = \lambda_1^k \left( c_1 v_1 + \frac{c_2}{\lambda_1} \frac{2 \lambda_1}{2} v_2 + \ldots + \frac{c_n}{\lambda_1} \frac{2 \lambda_1}{2} v_n \right)$$
On the last page we had
\[ A^h x_0 = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_n \lambda_n v_n = \sum_i \lambda_i (c_i v_i + \epsilon_i) \]
where \( \epsilon_i = c_2 \frac{\lambda_2}{\lambda_1} v_2 + \cdots + c_n \frac{\lambda_n}{\lambda_1} v_n. \)
Recall also that \( |\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \).
Thus \( \frac{\lambda_2}{\lambda_1} \to 0 \), \( \frac{\lambda_3}{\lambda_1} \to 0 \), \ldots, \( \frac{\lambda_n}{\lambda_1} \to 0 \).
and therefore \( \epsilon_i \to 0 \) as \( h \to \infty \).

Using this convergence we obtain
\[ \frac{A^h x_0}{\|A^h x_0\|} = \frac{\lambda_1 (c_1 \xi + \epsilon_1)}{\|\lambda_1 \| (c_1 \xi + \epsilon_1)} \to \pm \xi \] as \( h \to \infty \).
Recall that \( \|\xi\| = 1 \).

In other words:
\[ \frac{A^h x_0}{\|A^h x_0\|} \to \pm \xi \]
«eigenvector corresponding to largest eigenvalue \( \lambda_1 \)»

In order to find the eigenvalue \( \lambda_1 \), we just need to compute an inner product:
\[ \langle \xi_1, A \xi_1 \rangle = \lambda_1 \|\xi_1\|^2 = \lambda_1 \]
**ALGORITHM (Power method)**

**Input:** $A$, random $x_0 \in \mathcal{D}$, $N (= \text{max. iterations})$

**Iterate:** for $k = 1 : N$

\[ x_k = A \times x_{k-1} \]
\[ x_k = \frac{x_k}{\|x_k\|} \]

and

**Output:** $v_1 = x_N$
\[ \lambda_1 = \langle v_1, Av_1 \rangle \]

Remark: 1) In a serious algorithm we would not set a fixed number of maximal iterations but rather check for convergence, for example by checking if $\|x_k - x_{k-1}\|$ is smaller than some acceptable error and stopping if that is achieved.

2) So far the power method gives us only the eigenvalue with the longest absolute value, and its corresponding eigenvector. However, with a little modification (see next page) we can use this method to compute the eigenvalue that is closest to any number $s$. The key to this is that the eigenvalues of $(A - sI)^{-1}$ are exactly $(A_i - s)^{-1}$ as long as $(A - sI)$ is invertible. Moreover, the eigenvectors of $A$ and $(A - sI)^{-1}$ are the same.

Question: What happens if $(A - sI)$ is not invertible?
A modification of the power method

Let $s \in \mathbb{R}$ and suppose $A$ is $n \times n$ with real eigenvalues (as before). Then one of the following two holds:

- (a) $s$ is an eigenvalue of $A$, i.e. $A-sI$ is not invertible
- (b) The eigenvalues of $(A-sI)^{-1}$ are exactly $\frac{1}{f_j-s}$ and the corresponding eigenvectors are still $v_j$, $j=1, \ldots, n$.

**Proof:**

Let $Av = \lambda v$ for $v \neq 0$. Then, if $(A-sI)$ is invertible, we have

$$Av = (A-sI)v = \lambda v - sv = \lambda v - sv$$

$$\Rightarrow (A-sI)v = (\lambda - s)v$$

$$\Rightarrow (A-sI)^{-1}v = \frac{1}{\lambda - s}v$$

So $\frac{1}{\lambda - s}$ and $v$ are an eigenvalue/eigenvector pair for $(A-sI)^{-1}$. \qed

Hence using the power method on $(A-sI)^{-1}$ we obtain that the dominate eigenvalue of $(A-sI)^{-1}$ is $\frac{1}{f_j-s}$ if and only if $|\lambda_j-s| \leq |\lambda_x-s|$ for all $l \neq j$.

Thus the power method gives us the eigenvector $v_j$ corresponding to this $f_j$. Therefore to get the eigenvalue $f_j$ closest to $s$ we just compute $\langle v_j, Av_j \rangle$.

**Remark:** An example on how to implement the power method and the modified power method in MATLAB is given in Chapter 4 (page 167, eq. 170) in the text. \[x\]
Recursion relations

Example: \( a_n = a_{n-1} + 2 ; \ a_0 = 5 \) ← recursive definition of \( a_n \)

\( \Rightarrow 5, 7, 9, \ldots \)

Or we could find a formula:

\( a_n = a_0 + 2n \ ; \ a_0 = 5 \)

\( a_n = 5 + 2n \) ← direct formula for \( a_n \)

The Fibonacci sequence

\[ 0, 1, 1, 2, 3, 5, 8, 13, \ldots \]

(*) \( F_{n+1} = F_n + F_{n-1} \ ; \ F_0 = 0, \ F_1 = 1 \)

Recursion relation whose solution is the Fibonacci sequence.

Goal: Find a formula for \( F_n \).

How do we do that? \( \Rightarrow \) Rewrite the recursion (*) as a matrix equation.

(**) \[
\begin{bmatrix}
F_{n+1} \\
F_n
\end{bmatrix} =
\begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
F_n \\
F_{n-1}
\end{bmatrix} = V_n = A V_{n-1} \]

\[
\begin{bmatrix}
F_1 \\
F_0
\end{bmatrix} = \begin{bmatrix} 1 \\
0 \end{bmatrix} = V_0
\]

Note that (*) and (**) are equivalent!
Recall from last page: \[ v_0 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = v_n, \quad \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = v_{n-1}; \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \]

With this we have
\[ v_1 = A \cdot v_0; \quad v_2 = A^2 \cdot v_0; \quad \ldots \quad v_n = A^n \cdot v_0. \]

That is
\[ \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \]

Since \( A \) is real-symmetric (thus hermitian) we know that it is unitarily diagonalizable.

Eigenvalues of \( A \): \( \lambda_1 = \frac{1 + \sqrt{5}}{2} \); \( \lambda_2 = \frac{1 - \sqrt{5}}{2} \)

Exercise: verify this.

Corresponding eigenvectors: \( v_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} \); \( v_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} \)

We obtain:
\[ A = S \cdot D \cdot S^{-1} \quad \text{with} \quad S = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad S^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \lambda_2 \\ -1 & \lambda_1 \end{bmatrix} \]

and thus \( A^n = S \cdot D^n \cdot S^{-1} \quad \text{and so:} \)
\[ \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = S \cdot D^n \cdot S^{-1} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \ldots = \frac{1}{\sqrt{5}} \begin{bmatrix} \lambda_1^n & -\lambda_2^n \\ \lambda_1^n & \lambda_2^n \end{bmatrix} \Rightarrow F_n = \frac{1}{\sqrt{5}} \left( \lambda_1^n - \lambda_2^n \right) \]
On the last page we found:  \( F_n = \frac{1}{\sqrt{5}} (\phi^n - \phi_2^n) \).

Hence we have found a formula for \( F_n \). From this formula we get even more information of \( F_n \) for large \( n \):

Note that \( |\phi_2| = |\frac{1 - \sqrt{5}}{2}| < 1 \Rightarrow \phi_2^n \rightarrow 0 \) for \( n \rightarrow \infty \)

Thus for large \( n \) we have:  \( F_n \approx \frac{1}{\sqrt{5}} \phi_1^n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n \).

This is an integer up to a small error!

A more general example

Consider:  \( x_{n+1} = 3x_n + x_{n-1} + 2x_{n-2} \); \( x_0 = a \), \( x_1 = b \); \( x_2 = c \)

We want to "analyze" \( x_n \) for arbitrary \( n \in \mathbb{N} \), i.e. similar goal as before.

Solution:

\[
\begin{bmatrix}
x_{n+1} \\
x_n \\
x_{n-1}
\end{bmatrix} =
\begin{bmatrix}
3 & 2 & 0 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x_n \\
x_{n-1} \\
x_{n-2}
\end{bmatrix};
\]

\[ V_n =
\begin{bmatrix}
x_2 \\
x_1 \\
x_0
\end{bmatrix} =
\begin{bmatrix}
c \\
b \\
a
\end{bmatrix} \]

\[ \Rightarrow \quad V_n = A^n V_0 \]

\( \sim \) to "analyze" find eigenvalues/eigenvectors of \( A \), diagonalize \( A \), ...
The rest of this lecture is not relevant for the final!

However it is a nice application of the theory about recursion relations we have talked on today.
An example from quantum mechanics

We already know how to approximate differential equations by matrix equations. If we discretize the Schrödinger equation for an electron moving in a solid we obtain the so-called **tight binding model**.

Let us consider a single electron moving in a one dimensional semi-infinite crystal. We assume that the electron is constrained to live at discrete lattice points, labelled $0, 1, 2, \ldots$

For each lattice point $n$ there is a potential $V_n$ that describes how much the atom at that lattice point attracts or repels the electron. Note that positive $V_n$'s indicate repulsion and negative $V_n$'s indicate attraction.

Typical situations that are studied in physics are potentials $V_n$ that repeat the same pattern periodically ($\rightarrow$ crystals) or completely random potentials ($\rightarrow$ disordered media).

The so-called wavefunction for the electron is a sequence of complex numbers

$$\Psi = \{\psi_0, \psi_1, \psi_2, \ldots\}$$

The sequence $\Psi$ is called a bound state with energy $E$ if it satisfies the following three conditions:

1. The discret version of the time independent Schrödinger eq.: $-\psi_{n+1} - \psi_{n-1} + V_n \psi_n = E \psi_n$
2. The boundary condition: $\psi_0 = 0$
3. The normalization condition: $N^2 = \sum_{n=0}^{\infty} |\psi_n|^2 < \infty$
The conditions (1) - (3) are trivially satisfied if \( \Psi = \{0,0,0,\ldots\} \) so we specifically exclude this case. Moreover, for any given energy \( E \) we will always be able to find a wave function \( \Psi \) that satisfies condition (1) & (2). However, for most energies \( E \), none of the \( \Psi \)'s will be getting smaller for large \( n \), so the normalization condition (3) will not be satisfied.

It can be shown that there is only a discrete set of energies \( E \) for which a bound state exists, i.e., satisfies condition (1) - (3). In other words, the energy \( E \) is quantized.

Note that if \( E \) is one of the allowed energy values and \( \Psi \) is the corresponding bound state, then the numbers \( \frac{14n^2}{N^2} \) are interpreted as the probabilities of finding an electron with energy \( E \) at the \( n \)-th site.

Notice also that if \( \Psi \) is a bound state with energy \( E \), then so is any non-zero multiple \( a \Psi = \{a\Psi_0, a\Psi_1, a\Psi_2, \ldots\} \).

Now that we have defined the model we can use what we have learned about recursion relations to find the eigenvalues, i.e., energies \( E \) corresponding to bound states, of this model.

How this works will be presented on the next page.
The discrete Schrödinger equation (1) together with the initial condition (2) is a recursion relation.

\[
\begin{bmatrix}
\psi_{n+1} \\
\psi_n
\end{bmatrix} = \begin{bmatrix}
V_n - E & -1 \\
1 & 0
\end{bmatrix} \begin{bmatrix}
\psi_n \\
\psi_{n-1}
\end{bmatrix}
\]  

So if we set \( x_n = \begin{bmatrix} \psi_{n+1} \\ \psi_n \end{bmatrix} \) and define \( A(z) = \begin{bmatrix} z & -1 \\ 1 & 0 \end{bmatrix} \) then (*) implies

\[ x_n = A(V_n - E) A(V_{n-1} - E) \cdots A(V_1 - E) x_0 \]

Moreover condition (2) says that \( x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \) since \( \psi_0 = 0 \).

Recall that we can replace \( \Phi \) by \( a \psi_1 \) for any non-zero \( a \). Thus choosing \( a = \frac{1}{\psi_1} \) we can actually assume that

\[ x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

So for \( n \) we have computed \( x_n \) (and thus \( \psi_n \)) satisfying condition (1) & (2) for any values of \( V_1, V_2, \ldots \). Condition (3) is a statement about the large \( n \) behaviour of \( \psi_n \).

This can be very difficult to determine unless we know more about the values \( V_n \).

Thus to continue we need to consider the simplest possible situation:

\[ V_1 = -\alpha, \quad V_2 = V_3 = \ldots = 0. \]
Recall we now consider a potential with most values $\omega, 0: v_1 = -\alpha, v_2 = v_3 = \ldots = 0$.

In this situation we have

$$x_n = (A(-\omega))^n \cdot A(-\omega - \omega) x_0 = (A(-\omega))^n x_1,$$

where

$$x_1 = A(-\omega - \omega) x_0 = \begin{bmatrix} -\omega & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\omega \omega + \omega \end{bmatrix}$$

The large $n$ behaviour of $x_n$ can be computed using the eigenvalues and eigenvectors of $A(-\omega)$. Suppose they are $\lambda_1, v_1$ and $\lambda_2, v_2$. Then we expand

$$x_1 = a_1 v_1 + a_2 v_2$$

and conclude that

$$x_n = (A(-\omega))^n (a_1 v_1 + a_2 v_2) = a_1 \lambda_1^n v_1 + a_2 \lambda_2^n v_2$$

Note that all quantities in this equation depend on $\omega$ and our goal is to choose $\omega$ such that $x_n$ becomes small for large $n$. Moreover note that $\det (A(-\omega)) = 1 \Rightarrow \lambda_1 \lambda_2 = 1$

**Case 1: (complex eigenvalues)** Since $A(-\omega)$ has real entries, the eigenvalues must be complex conjugates of each other in this case. Thus $\lambda_2 = \overline{\lambda_1}$ and therefore $1 = \lambda_1 \lambda_2 = \lambda_1 \overline{\lambda_1} = 1$.

This means $\lambda_1$ and $\lambda_2$ lie on the unit circle in the complex plane, i.e. $\lambda_1 = e^{i\theta}, \lambda_2 = e^{-i\theta}$ for some $\theta$.

This implies that $\lambda_1^{-n} = e^{in\theta}$ is also on the unit circle and is not getting small for large $n$. Similarly for $\lambda_2$. So complex eigenvalues will never lead to bound states.

In fact complex eigenvalues correspond to scattering states and energy values corresponding to these states are energies at which the electron can move freely through the crystal.
Case 2: (real eigenvalues)

If \(|\lambda_1| > 1\) then \(1/|\lambda_1| < 1\) and vice versa. So one of the products \(\lambda^{n-1}_1, \lambda^{n-1}_2\) will be growing larger, and one will be getting smaller. So the only way that \(x_n\) can be getting small is if the coefficient \(a_1\) or \(a_2\) sitting in front of the growing product is \(2\pi, 0\).

Let us calculate the eigenvalues of \(A(-E)\), i.e. the roots of \(\det (A(-E) - \lambda I) = 0\).
We obtain
\[
\lambda = \frac{-E \pm \sqrt{E^2 - 4}}{2}.
\]

If \(-2 < E < 2\) the eigenvalues are complex, so there are no bound states. The interval \([-2, 2]\) is the conduction band where the electrons can move through the crystal.

Note that we included here \(E = \pm 2\) for which there is only one eigenvalue, namely \(1\). In this case there is only one eigenvector and our analysis does not apply. However, it can be shown that there are no bound states.

\(E < -2:\) The largest eigenvalue is \(\lambda_1 = \frac{-E + \sqrt{E^2 - 4}}{2}\) with corresponding eigenvector \(v_1 = \begin{bmatrix} -1 \\ E + \lambda_1 \end{bmatrix}\).

\(E > 2:\) The smallest eigenvalue is \(\lambda_2 = \frac{-E - \sqrt{E^2 - 4}}{2}\) with eigenvector \(v_2 = \begin{bmatrix} -1 \\ E + \lambda_2 \end{bmatrix}\).

We must now compute \(a_1, a_2\) and determine when it is \(\pi, 0\).

We have \([v_1 v_2] [a_1, a_2] = x_1 \Rightarrow a_1 = (\lambda_1, -\lambda_2)^{-1} \left((-a + E)(E + \lambda_2) + 1\right) \Rightarrow a_1 = 0 \Rightarrow (a + E)(E - \sqrt{E^2 - 4}) - 2 = 0\)

Example: \(a = 5 \Rightarrow E = -5.2\) and thus we have if \(v_1 = 5, v_2 = v_3 = \ldots = 0\), exactly one bound state with energy \(E = -5.2\). In the diagram of energy spectrum, \(-6 -4 -2 0 2 \rightarrow E\) indicates the conduction band and the bound state is at \(-5.2\).