Chapter 4: Eigenvalue Problems

Eigenvalues and Eigenvectors

- Standard *eigenvalue problem: Given $n \times n$ matrix $\mathbf{A}$, find scalar λ and nonzero vector $\mathbf{x}$ such that*

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

- λ is the *eigenvalue* and $\mathbf{x}$ is the *eigenvector*

- λ (and $\mathbf{x}$) may be complex even if $\mathbf{A}$ is real

- *Spectrum* of $\mathbf{A}$ = set of all eigenvalues $\lambda(\mathbf{A})$

- *Spectral radius* $\rho(\mathbf{A}) = \max\{|\lambda| : \lambda \in \lambda(\mathbf{A})\}$

Geometric Interpretation

- The matrix-vector product $\hat{\mathbf{v}} = A\mathbf{v}$ stretches or shrinks any vector $\mathbf{v}$ lying in direction of eigenvector $\mathbf{x}$

- Scalar expansion or contraction factor is given by corresponding λ

- Eigenvalues and eigenvectors lead to simple interpretation of general linear transformations (e.g., as represented by matrix-vector products)

- They are particularly useful when considering iterative processes that can be cast as a sequence of matrix-vector products, such as

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0, \;\; \mathbf{x}_2 = \mathbf{A}\mathbf{x}_1, \; \dots \; , \mathbf{x}_k = \mathbf{A}^k\mathbf{x}_0,$$

- Such sequences are in fact at the core of most of the algorithms used to find the eigenpairs $(\lambda, \mathbf{x})$ of $\mathbf{A}$

Examples: Eigenvalues and Eigenvectors

- $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$: $\quad \lambda_1 = 1,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \lambda_2 = 2,\ \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$: $\quad \lambda_1 = 1,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \lambda_2 = 2,\ \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$: $\quad \lambda_1 = 2,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = 4,\ \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$: $\quad \lambda_1 = 2,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = 1,\ \mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$: $\quad \lambda_1 = i,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \lambda_2 = -i,\ \mathbf{x}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$

Examples: Eigenvalues and Eigenvectors

- $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$: $\quad \lambda_1 = 1,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \lambda_2 = 2,\ \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$: $\quad \lambda_1 = 1,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \lambda_2 = 2,\ \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$: $\quad \lambda_1 = 2,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = 4,\ \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Examples: Eigenvalues and Eigenvectors

- $\mathbf{A} = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$: $\quad \lambda_1 = 2,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = 1,\ \mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

- $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$: $\quad \lambda_1 = i,\ \mathbf{x}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \lambda_2 = -i,\ \mathbf{x}_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$

Classic Eigenvalue Problem

- Consider the coupled pair of differential equations:

$$\begin{aligned}\frac{dv}{dt} &= 4v - 5w, \quad v = 8 \text{ at } t = 0,\\ \frac{dw}{dt} &= 2v - 3w, \quad w = 5 \text{ at } t = 0.\end{aligned}$$

- This is an ***initial-value problem***.
- With the coefficient matrix,

$$A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix},$$

we can write this as,

$$\frac{d}{dt}\begin{pmatrix} v(t) \\ w(t) \end{pmatrix} = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}\begin{pmatrix} v(t) \\ w(t) \end{pmatrix}.$$

- Introducing the *vector unknown*, $\mathbf{u}(t) := [v(t)\ w(t)]^T$ with $\mathbf{u}(0) = [8\ 5]^T$, we can write the system in vector form,

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}, \quad \text{with } \mathbf{u} = \mathbf{u}(0) \text{ at } t = 0.$$

- How do we find $\mathbf{u}(t)$?

- If we had a 1×1 matrix $A = a$, we would have a scalar equation:

$$\frac{du}{dt} = a\,u \;\text{ with } u = u(0) \text{ at } t = 0.$$

  The solution to this equation is a pure exponential:

$$u(t) = e^{at}\,u(0),$$

  which satisfies the initial condition because $e^0 = 1$.

- The derivative with respect to t is $ae^{at}u(0) = au$, so it satisfies the scalar initial value problem.

- The constant a is critical to how this system behaves.
  - If $a > 0$ then the solution grows in time.
  - If $a < 0$ then the solution decays.
  - If $a \in Im$ then the solution is oscillatory.
    (More on this later...)

- Coming back to our system, suppose we again look for solutions that are pure exponentials in time, e.g.,

$$\begin{aligned} v(t) &= e^{\lambda t} y \\ w(t) &= e^{\lambda t} z. \end{aligned}$$

- If this is to be a solution to our initial value problem, we require

$$\begin{aligned} \frac{dv}{dt} &= \lambda e^{\lambda t} y = 4e^{\lambda t} y - 5e^{\lambda t} z \\ \frac{dw}{dt} &= \lambda e^{\lambda t} z = 2e^{\lambda t} y - 3e^{\lambda t} z. \end{aligned}$$

- The $e^{\lambda t}$ cancels out from each side, leaving:

$$\begin{aligned} \lambda y &= 4y - 5z \\ \lambda z &= 2y - 3z, \end{aligned}$$

which is the eigenvalue problem.

$$\begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

- In vector form, $\mathbf{u}(t) = e^{\lambda t}\mathbf{x}$, yields

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} \iff \lambda e^{\lambda t}\mathbf{x} = A(e^{\lambda t}\mathbf{x})$$

  which gives the eigenvalue problem in matrix form:

$$\lambda\mathbf{x} = A\mathbf{x} \quad \text{or}$$

$$A\mathbf{x} = \lambda\mathbf{x}.$$

- As in the scalar case, the solution behavior depends on whether λ has
  - positive real part $\longrightarrow$ a growing solution,
  - negative real part $\longrightarrow$ a decaying solution,
  - an imaginary part $\longrightarrow$ an oscillating solution.
- Note that here we have two unknowns: λ and $\mathbf{x}$.
- We refer to $(\lambda, \mathbf{x})$ as an eigenpair, with *eigenvalue* λ and *eigenvector* $\mathbf{x}$.

Solving the Eigenvalue Problem

- The eigenpair satisfies

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

which is to say,

  - $\mathbf{x}$ is in the null-space of $A - \lambda I$
  - λ is chosen so that $A - \lambda I$ has a null-space.

- We thus seek λ such that $A - \lambda I$ is singular.
- Singularity implies $det(A - \lambda I)$=0.
- For our example:

$$0 = \begin{vmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{vmatrix} = (4-\lambda)(-3-\lambda) - (-5)(2),$$

or

$$\lambda^2 - \lambda - 2 = 0,$$

which has roots $\lambda = -1$ or $\lambda = 2$.

Finding the Eigenvectors

- For the case $\lambda = \lambda_1 = -1$, $(A - \lambda_1 I)\mathbf{x}_1$ satisfies,

$$\begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which gives us the eigenvector $\mathbf{x}_1$

$$\mathbf{x}_1 = \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- **Note that any nonzero multiple of $\mathbf{x}_1$ is also an eigenvector.**
- Thus, $\mathbf{x}_1$ defines a *subspace* that is invariant under multiplication by A.

Because $Ax_1 = \lambda\, x_1$, i.e., it is simply a stretching of x_1

- For the case $\lambda = \lambda_2 = 2$, $(A - \lambda_2 I)\mathbf{x}_2$ satisfies,

$$\begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

which gives us the second eigenvector as any multiple of

$$\mathbf{x}_2 = \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}.$$

Return to Model Problem

- Note that our model problem $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$, is *linear* in the unknown $\mathbf{u}$.
- Thus, if we have two solutions $\mathbf{u}_1(t)$ and $\mathbf{u}_2(t)$ satisfying the differential equation, their sum $\mathbf{u} := \mathbf{u}_1 + \mathbf{u}_2$ also satisfies the equation:

$$
\begin{array}{rl}
 & \dfrac{d\mathbf{u}_1}{dt} = A\mathbf{u}_1 \\[2ex]
+ & \dfrac{d\mathbf{u}_2}{dt} = A\mathbf{u}_2 \\
\hline
 & \dfrac{d}{dt}(\mathbf{u}_1 + \mathbf{u}_2) = A(\mathbf{u}_1 + \mathbf{u}_2) \\[2ex]
 & \dfrac{d\mathbf{u}}{dt} = A\mathbf{u}
\end{array}
$$

- Take $\mathbf{u}_1 = c_1 e^{\lambda_1 t}\mathbf{x}_1$:

$$\begin{aligned}\frac{d\mathbf{u}_1}{dt} &= c_1\lambda_1 e^{\lambda_1 t}\mathbf{x}_1 \\ A\mathbf{u}_1 &= A\left(c_1 e^{\lambda_1 t}\mathbf{x}_1\right) \\ &= c_1 e^{\lambda_1 t} A\mathbf{x}_1 \\ &= c_1 e^{\lambda_1 t}\lambda_1\mathbf{x}_1 \\ &= \frac{d\mathbf{u}_1}{dt}.\end{aligned}$$

- Similarly, for $\mathbf{u}_2 = c_2 e^{\lambda_2 t}\mathbf{x}_2$:

$$\frac{d\mathbf{u}_2}{dt} = A\mathbf{u}_2.$$

- Thus,

$$\begin{aligned}\frac{d\mathbf{u}}{dt} &= \frac{d}{dt}\left(\mathbf{u}_1 + \mathbf{u}_2\right) = A\left(\mathbf{u}_1 + \mathbf{u}_2\right) \\ \mathbf{u} &= c_1 e^{\lambda_1 t}\mathbf{x}_1 + c_2 e^{\lambda_2 t}\mathbf{x}_2.\end{aligned}$$

- The only remaining part is to find the coefficients c_1 and c_2 such that $\mathbf{u} = \mathbf{u}(0)$ at time $t = 0$.

- This initial condition yields a 2×2 system,

$$\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \end{pmatrix}.$$

- Solving for c_1 and c_2 via Gaussian elimination:

$$\begin{bmatrix} 1 & 5 \\ 1 & 2 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \end{pmatrix}$$

$$\begin{bmatrix} 1 & 5 \\ 0 & -3 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 8 \\ -3 \end{pmatrix}$$

$$\begin{aligned} c_2 &= 1 \\ c_1 &= 8 - 5c_1 = 3. \end{aligned}$$

- So, our solution is
$$\begin{aligned} \mathbf{u}(t) &= \mathbf{x}_1 c_1 e^{\lambda_1 t} + \mathbf{x}_2 c_2 e^{\lambda_2 t} \\ &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} 3e^{-t} + \begin{pmatrix} 5 \\ 2 \end{pmatrix} e^{2t}. \end{aligned}$$

- Clearly, after a long time, the solution is going to look like a multiple of $\mathbf{x}_2 = [5\ 2]^T$ because the component of the solution parallel to $\mathbf{x}_1$ will decay.

- (More precisely, the component parallel to $\mathbf{x}_1$ will not grow as fast as the component parallel to $\mathbf{x}_2$.)

Example Summary

- Model problem, $\mathbf{u} \in \mathcal{R}^n$,

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}, \quad \mathbf{u} = \mathbf{u}(0) \text{ at time } t = 0.$$

- Assuming A has n *linearly independent* eigenvectors, can express

$$\mathbf{u}(t) = \sum_{j=1}^{n} \mathbf{x}_j c_j e^{\lambda_j t}.$$

- Coefficients c_j determined by initial condition:

$$X\mathbf{c} = \sum_{j=1}^{n} \mathbf{x}_j c_j = \mathbf{u}(0) \iff \mathbf{c} = X^{-1}\mathbf{u}(0).$$

- Eigenpairs $(\lambda_j, \mathbf{x}_j)$ satisfy

$$A\mathbf{x}_j = \lambda_j \mathbf{x}_j.$$

Growing / Decaying Modes

- Our model problem,

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} \longrightarrow \mathbf{u}(t) = \mathbf{x}_1 c_1 e^{\lambda_1 t} + \mathbf{x}_2 c_2 e^{\lambda_2 t}$$

  leads to growth/decay of components.

- Also get growth/decay through matrix-vector products.
- Consider $\mathbf{u} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2$.

$$\begin{aligned} A\mathbf{u} &= c_1 A\mathbf{x}_1 + c_2 A\mathbf{x}_2 \\ &= c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 \end{aligned}$$

$$\begin{aligned} A^k\mathbf{u} &= c_1\lambda_1^k\mathbf{x}_1 + c_2\lambda_2^k\mathbf{x}_2 \\ &= \lambda_2^k \left[c_1 \left(\frac{\lambda_1}{\lambda_2}\right)^k \mathbf{x}_1 + c_2\mathbf{x}_2 \right]. \end{aligned}$$

$$\lim_{k\longrightarrow\infty} A^k\mathbf{u} = \lambda_2^k \left[c_1 \cdot 0 \cdot \mathbf{x}_1 + c_2\mathbf{x}_2\right] = c_2\lambda_2^k\mathbf{x}_2.$$

- So, repeated matrix-vector products lead to emergence of eigenvector associated with the eigenvalue λ that has largest modulus.
- This is the main idea behind the ***power method***, which is a common way to find the eigenvector associated with $\max|\lambda|$.

Characteristic Polynomial

- Equation $\mathbf{Ax} = \lambda\mathbf{x}$ is equivalent to

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

  which has nonzero solution $\mathbf{x}$ *iff* matrix $(\mathbf{A} - \lambda\mathbf{I})$ is singular

- Eigenvalues of $\mathbf{A}$ are roots λ_i of *characteristic polynomial*

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

  of degree n in λ

- *Fundamental Theorem of Algebra* implies that $n \times n$ matrix $\mathbf{A}$ always has n eigenvalues, but they need not be real nor distinct

- Complex eigenvalues of real matrix occur in complex conjugate pairs: If $\lambda = \alpha + i\beta$ is an eigenvalue of a real matrix then so is $\alpha - i\beta$, where $i = \sqrt{-1}$

Example: Characteristic Polynomial

- Evaluate $\det(\mathbf{A} - \lambda\mathbf{I})$ of earlier example

$$
\begin{aligned}
|\mathbf{A} - \lambda\mathbf{I}| &= \det\left(\begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{bmatrix}\right) \\
&= (3-\lambda)(3-\lambda) - (-1)(-1) = \lambda^2 - 6\lambda + 8 = 0
\end{aligned}
$$

- Eigenvalues are

$$
\lambda = \frac{6 \pm \sqrt{36-32}}{2}, \quad \text{or} \quad \lambda_1 = 2, \quad \lambda_2 = 4
$$

Companion Matrix

- Monic polynomial

$$p(\lambda) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + c_n\lambda^n$$

is characteristic polynomial of *companion matrix*

$$\mathbf{C}_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{bmatrix}$$

- Roots of polynomial degree > 4 cannot always be computed in finite number of steps

- So in general, computation of eigenvalues of matrices of order > 4 requires a (theoretically infinite) iterative process

Example: Companion Matrix, $n = 3$

- Consider companion matrix

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & -c_0 \\ 1 & 0 & -c_1 \\ 0 & 1 & -c_2 \end{bmatrix}$$

- Evaluate determinant of $\mathbf{C} - \lambda \mathbf{I}$

$$|\mathbf{C} - \lambda \mathbf{I}| = \begin{vmatrix} -\lambda & 0 & -c_0 \\ 1 & -\lambda & -c_1 \\ 0 & 1 & -(c_2 + \lambda) \end{vmatrix}$$

$$= -c_0 \begin{vmatrix} 1 & -\lambda \\ 0 & 1 \end{vmatrix} + c_1 \begin{vmatrix} -\lambda & 0 \\ 0 & 1 \end{vmatrix} - (c_2 + \lambda) \begin{vmatrix} -\lambda & 0 \\ 1 & -\lambda \end{vmatrix}$$

$$= -c_0 - c_1 \lambda - c_2 \lambda^2 - \lambda^3 = 0$$

- Roots of resultant monic polynomial, $p(\lambda) = c_0 + c_1 \lambda + c_2 \lambda^2 + \lambda^3 = 0$, are the 3 eigenvalues, λ_1, λ_2, and λ_3

Characteristic Polynomial, continued

- Computing eigenvalues using characteristic polynomial is *not recommended* because of
  - work in computing coefficients of characteristic polynomial
  - sensitivity of coefficients of characteristic polynomial
  - work in solving for roots of characteristic polynomial
- Characteristic polynomial is a powerful theoretical tool but usually not useful computationally
- In fact, in many cases we use eigenvalue solvers to find the roots of poynomials

Example: Characteristic Polynomial

- Consider $\mathbf{A} = \begin{bmatrix} 1 & \epsilon \\ \epsilon & 1 \end{bmatrix}$

  with $\epsilon_M < \epsilon < \sqrt{\epsilon_M}$

- Exact eigenvalues of $\mathbf{A}$ are $1 + \epsilon$ and $1 - \epsilon$

- Computing characteristic polynomial in float point arithmetic leads to

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^2 - 2\lambda + (1 - \epsilon^2) = \lambda^2 - 2\lambda + 1$$

  which as 1 as a double root

- Thus, eigenvalues cannot be resolved by this method even though they are distinct to working precision

Multiplicity and Diagonalizability

- *Multiplicity* is number of times root appears when polynomial is written as product of linear factors

Algebraic Multiplicity

- Eigenvalue with multiplicity 1 is *simple*

- *Defective* matrix has eigenvalue of $k > 1$ with fewer than k linearly independ corresponding eigenvectors

Geometric Multiplicity

- Nondefective matrix $\mathbf{A}$ has n linearly independent eigenvectors, so it is *diagonalizable*

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{D}$$

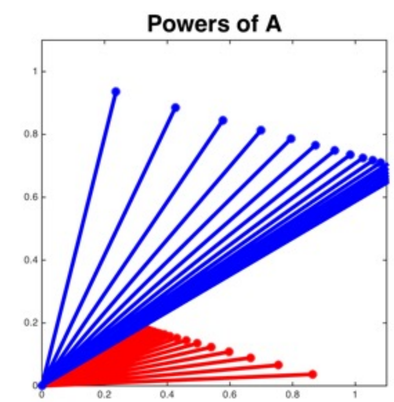
  where $\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2\ \cdots \mathbf{x}_n]$ is nonsingular matrix of eigenvectors

- Note: every matrix is ϵ away from being diagonalizable

Diagonalization

- The real merit of eigenvalue decomposition is that it simplifies powers of a matrix.
- Consider

$$\begin{aligned}
X^{-1}AX &= D, \ \textit{diagonal} \\
AX &= XD \\
A &= XDX^{-1} \\
A^2 &= \left(XDX^{-1}\right)\left(XDX^{-1}\right) \\
&= XD^2X^{-1} \\
A^k &= \left(XDX^{-1}\right)\left(XDX^{-1}\right)\cdots\left(XDX^{-1}\right) \\
&= XD^kX^{-1} \\
&= X\begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}X^{-1}
\end{aligned}$$



pow_a.m

- High powers of A tend to be dominated by largest eigenpair $(\lambda_1, \underline{x}_1)$, assuming $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$.

Matrix Powers Example

- Consider our 1D finite difference example introduced earlier.

$$\frac{d^2u}{dx^2} \;=\; f(x) \;\longrightarrow\; -\frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} \;\approx\; f(x_i).$$

  where $u(0) = u(1) = 0$ and $\Delta x = 1/(n+1)$.

- In matrix form,

$$A\mathbf{u} = \frac{1}{\Delta x^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_m \end{pmatrix}$$

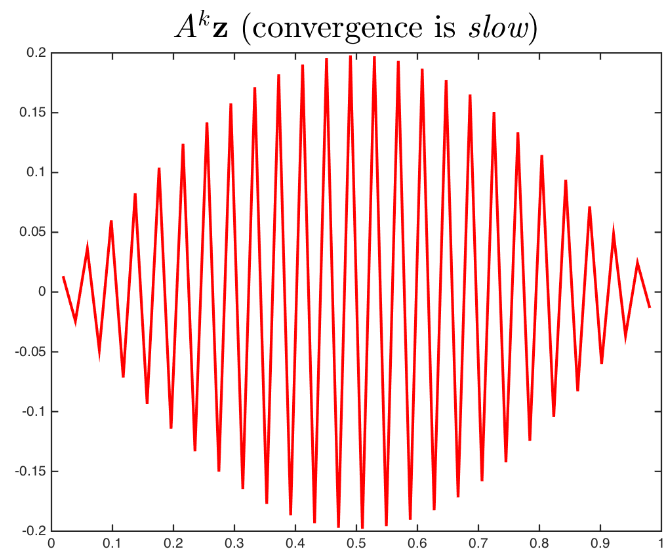
- Eigenvectors and eigenvalues have closed-form expression:

$$(\mathbf{z}_k)_i \;=\; \sin k\pi x_i \;=\; \sin k\pi i \Delta x \qquad\qquad \lambda_k \;=\; \frac{2}{\Delta x^2}\left(1 - \cos k\pi \Delta x\right)$$

- Eigenvalues are in the interval $\sim\; [\pi^2,\; 4(n+1)^2]$.

Matlab Example: heat_demo.m

- Repeatedly applying A to a random input vector reveals the eigenvalue of maximum modulus.
- This idea leads to one of the most common (but not most efficient) ways of finding an eigenvalue/vector pair, called the power method.




```
hdr

n = 50;
h = 1/(n+1);
e = ones(n,1);
A = spdiags([-e 2*e -e],-1:1, n,n)/(h*h);
x = 1:n; x=h*x';

z=sin(pi*x);
for k=1:3000;
    z=A*z; z=z/norm(z);
    plot(x,z,'r-',lw,2);
    title('$A^k {\bf z}$ (convergence is {\em slow})',intp,ltx,fs,24);
    drawnow;
    k
end;
```

```
hdr

n=100;


h = 1/(n+1); e = ones(n,1); A = spdiags([-e 2*e -e],-1:1, n,n)/(h*h);

x=1:n; x=h*x';

z=rand(n,1); hold off;
z=sin(pi*x);
for k=1:2000;
    z=A*z; z=z/norm(z);
    plot(x,z,'r-',lw,2); pause;
    k
end;
if -min(z) > max(z); z=-z; end; plot(x,z,'r-'); pause;

[Z,D]=eig(full(A)); zn=Z(:,n);


hold on
zn=zn/norm(zn); if -min(zn) > max(zn); zn=-zn; end;
plot(x,zn,'kx')
sn = sin(n*pi*x); sn=sn/norm(sn);
plot(x,sn,'go')
```

Diagonalization

- Note that if we define $A^0 = I$, we have any polynomial of A defined as

$$p_k(A)\underline{x} \;=\; X \begin{bmatrix} p_k(\lambda_1) & & & \\ & p_k(\lambda_2) & & \\ & & \ddots & \\ & & & p_k(\lambda_n) \end{bmatrix} X^{-1}\underline{x}.$$

- We can further extend this to other functions,

$$f(A)\underline{x} \;=\; X \begin{bmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \ddots & \\ & & & f(\lambda_n) \end{bmatrix} X^{-1}\underline{x}.$$

- For example, the solution to $f(A)\underline{x} = \underline{b}$ is would be

$$\underline{x} = X\left[f(D)\right]^{-1} X^{-1}\underline{b}.$$

- The diagonalization concept is very powerful because it transforms *systems* of equations into scalar equations.

Eigenspaces and Invariant Subspaces

- Eigenvectors can be scaled arbitrarily: if $\mathbf{Ax} = \lambda\mathbf{x}$, then $\mathbf{A}(\gamma\mathbf{x}) = \lambda(\gamma\mathbf{x})$ for any scalar γ, so $\gamma\mathbf{x}$ is also eigenvector corresponding to λ

- Eigenvectors are usually *normalized* by requiring some norm of eigenvector to be 1 (2-norm is most favored...)

- *Eigenspace* $= \mathcal{S}_\lambda = \{\mathbf{x} : \mathbf{Ax} = \lambda\mathbf{x}\}$

- Subspace $\mathcal{S}$ of $\mathbb{R}^n$ (or $\mathbb{C}^n$) is *invariant* if $\mathbf{A}\mathcal{S} \subseteq \mathcal{S}$

- For eigenvectors $\mathbf{x}_1 \cdots \mathbf{x}_p$ span$([\mathbf{x}_1 \cdots \mathbf{x}_p])$ is invariant subspace

- **Q:** When might invariance fail?

  A: In floating-point arithmetic, because of round-off error

Relevant Properties of Matrices

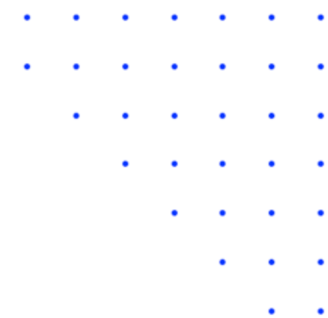
- Properties of matrix $\mathbf{A}$ relevant to eigenvalue problems

Property	Definition
diagonal	$a_{ij} = 0$ for $i \neq j$
tridiagonal	$a_{ij} = 0$ for $\|i - j\| > 1$
triangular	$a_{ij} = 0$ for $i > j$ (upper)
	$a_{ij} = 0$ for $i < j$ (lower)
Hessenberg	$a_{ij} = 0$ for $i > j + 1$ (upper)
	$a_{ij} = 0$ for $i < j - 1$ (lower)
orthogonal	$\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T = \mathbf{I}$
unitary	$\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H = \mathbf{I}$ $(\mathbf{A} \in \mathbb{C}^{ntimesn}$
symmetric	$\mathbf{A} = \mathbf{A}^T$
skew-symmetric	$\mathbf{A} = -\mathbf{A}^T$
Hermitian	$\mathbf{A} = \mathbf{A}^H$
normal	$\mathbf{A} = \mathbf{A}^H$
normal	$\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$

Upper Hessenberg (from last lecture…)

- A is upper Hessenberg – *A is upper triangular with one additional nonzero diagonal below the main one:* $A_{ij} = 0$ *if* $i > j+1$

0.1967	0.2973	0.0899	0.3381	0.5261	0.3965	0.1279
0.0934	0.0620	0.0809	0.2940	0.7297	0.0616	0.5495
0	0.2982	0.7772	0.7463	0.7073	0.7802	0.4852
0	0	0.9051	0.0103	0.7814	0.3376	0.8905
0	0	0	0.0484	0.2880	0.6079	0.7990
0	0	0	0	0.6925	0.7413	0.7343
0	0	0	0	0	0.1048	0.0513



- Requires only n Givens rotations, instead of $O(n^2)$, to effect QR factorization.

Examples: Matrix Properties

- Transpose: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$
- Conjugate transpose: $\begin{bmatrix} 1+i & 1+2i \\ 2-i & 2-2i \end{bmatrix}^H = \begin{bmatrix} 1-i & 2+i \\ 1-2i & 2+2i \end{bmatrix}$
- Symmetric: $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ Skew-Symmetric: $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}^T$
- Nonsymmetric: $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$
- Hermitian: $\begin{bmatrix} 1 & 1+i \\ 1-i & 2 \end{bmatrix}$
- NonHermitian: $\begin{bmatrix} 1 & 1+i \\ 1+i & 2 \end{bmatrix}$

Examples, continued

- Orthogonal: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$
- Unitary: $\begin{bmatrix} i\sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & -i\sqrt{2}/2 \end{bmatrix}$
- Nonorthogonal: $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$
- Normal: $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}$
- Nonnormal: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ ← *"canonical non-normal matrix"*
  Defective – has only one eigenvector.

Normal Matrices

Normal matrices have orthogonal eigenvectors, so $\underline{x}_i^H \underline{x}_j = \delta_{ij}$

Beware!

- If A is normal, it *has* orthogonal eigenvectors.
- That does not mean that all eigensolvers will *return* orthogonal eigenvectors.
- In particular, if two or more eigenvectors share the same eigenvalue, then they needn't be orthogonal to each other.
- You probably need to orthogonalize them yourself.

Normal ma

- symme
- skew-symmetric $(A = -A^T)$
- unitary $(U^H U = I)$
- circulant (periodic+Toeplitz)
- others ...

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

Properties of Eigenvalue Problems

Properties of eigenvalue problem affecting choice of algorithm and software

- Are all eigenvalues needed, or only a few?
- Are only eigenvalues needed, or are corresponding eigenvectors also needed?
- Is the matrix real or complex?
- Is the matrix relatively small and dense, or large and sparse?
- Does the matrix have any special properties such as symmetry, or is it a general matrix?

Sparsity

- Sparsity, either direct or implied, is a big driver in choice of eigenvalue solvers.
- Typically, only O(n) entries in entire matrix, where n ~ 10^9—10^{18} might be anticipated.
- Examples include Big Data (e.g., google page rank) and physics simulations (fluid, heat transfer, electromagnetics, fusion, etc.).
- Usually, need only a few (k << n) eigenvectors / eigenvalues.
- Often, there are special properties of A that make it difficult to create A. Instead, work strictly with matrix-vector products
- $\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}$

Conditioning of Eigenvalue Problems

- Condition of eigenvalue problem is sensitivity of eigenvalues and eigenvectors to changes in matrix

- Condition of eigenvalue problem is *not* same as conditioning of solution to linear system for same matrix

  - Finding $\lambda = 0$ is a common situation in eigenvalue problems, but indicates a singularity when trying to solve $\mathbf{Ax} = \mathbf{b}$ sensitivity of coefficients of characteristic polynomial

- Different eigenvalues and eigenvectors are not necessarily equally sensitive to perturbations in matrix

Conditioning of Eigenvalues

- If μ is eigenvalue of $\mathbf{A} + \mathbf{E}$ of nondefective matrix $\mathbf{A}$, then

  $$|\mu - \lambda_k| \leq \text{cond}_2(\mathbf{X}) \|\mathbf{E}\|_2$$

  where λ_k is closest eigenvalue of $\mathbf{A}$ to μ and $\mathbf{X}$ is the nonsingular matrix of eigenvectors of $\mathbf{A}$

- Absolute condition number of eigenvalues is condition number of matrix of eigenvectors with respect so solving linear equations (e.g., $\mathbf{X}\mathbf{c} = \mathbf{b}$)

- Eigenvalues may be sensitive if eigenvectors are nearly linearly dependent (i.e., matrix is nearly defective)

- For *normal* matrix ($\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$), eigenvectors are orthogonal, so eigenvalues are well-conditioned

Conditioning of Eigenvalues

- If $(\mathbf{A}+\mathbf{E})(\mathbf{x}+\Delta\mathbf{x}) = (\lambda+\Delta\Lambda)(\mathbf{x}+\Delta\mathbf{x})$, where λ is a simple eigenvalue of $\mathbf{A}$, then

$$|\Delta\lambda| \;\lesssim\; \frac{\|\mathbf{y}\|_2 \cdot \|\mathbf{x}\|_2}{|\mathbf{y}^H\mathbf{x}|}\|\mathbf{E}\|_2 \;=\; \frac{1}{\cos\theta}\|\mathbf{E}\|_2$$

  where $\mathbf{x}$ and $\mathbf{y}$ are corresponding right and left eigenvectors and θ is the angle between them

- For symmetric or Hermitian matrix right and left eigenvectors aare same so $\cos\theta = 1$ and eigenvalues are inherently well-conditioned

- Eigenvalues of nonnormal matrices may be sensitive

- For multiple or closely clustered eigenvalues, corresponding eigenvectors may be sensitive

Problem Transformations

- *Shift*: If $\mathbf{Ax} = \lambda\mathbf{x}$ and σ is any scalar, then $(\mathbf{A} - \sigma\mathbf{I})\mathbf{x} = (\lambda - \sigma)\mathbf{x}$, so eigenvalues of shifted matrix are shifted eigenvalues of $\mathbf{A}$

- *Inversion*: If $\mathbf{A}$ is nonsingular and $\mathbf{Ax} = \lambda\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$, then $\lambda \neq 0$ and $\mathbf{A}^{-1}\mathbf{x} = (1/\lambda)\mathbf{x}$, so eigenvalues of inverse are reciprocals of $\lambda(\mathbf{A})$

- *Powers*: If $\mathbf{Ax} = \lambda\mathbf{x}$, then $\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}$, so eigenvalues of power of matrix are λ^k

- *Polynomial*: If $\mathbf{Ax} = \lambda\mathbf{x}$, and $p(t)$ is a polynomial, then $p(\mathbf{A})\mathbf{x} = p(\lambda)\mathbf{x}$, so eigenvalues of polynomial in $\mathbf{A}$ are $p(\lambda)$.

Similarity Transformation

- $\mathbf{B}$ is *similar* to $\mathbf{A}$ if there exists a nonsingular matrix $\mathbf{T}$ such that

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

- Then,

$$\mathbf{B}\mathbf{y} = \lambda\mathbf{y} \implies \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{y} = \lambda\mathbf{y} \implies \mathbf{A}\mathbf{T}\mathbf{y} = \lambda\mathbf{T}\mathbf{y}$$

  so $\mathbf{A}$ and $\mathbf{B}$ have the same eigenvalues, and if $\mathbf{y}$ is eigenvector of $\mathbf{B}$, then $\mathbf{x} = \mathbf{T}\mathbf{y}$ is eigenvector of $\mathbf{A}$

- Similarity transformations preserve eigenvalues and eigenvectors are easily recovered

Example: Similarity Transformation

- From eigenvalues and eigenvectors for previous example,

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

  and hence

$$\begin{bmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

- So original matrix is similar to diagonal matrix, and eigenvectors form columns of similarity transformation matrix

Diagonal Form

- Eigenvalues of diagonal matrix are diagonal entries, eigenvectors are $\mathbf{X} = \mathbf{I}$
- Diagonal form is desirable in simplifying eigenvalue problems for general matrices by similarity transformations
- But not all matrices are diagonalizable
- Closest one can get, in general, is Jordan form, which is nearly diagonal but may have som nonzero entries on first superdiagonal corresponging to one or more multiple eigenvalues

Simple non-diagonalizable example, 2 x 2 Jordan block:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = 0$$

Only one eigenvector: $\underline{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

3×3 Non-Diagonalizable Example

$$A = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

- Characteristic polynomial is $(\lambda - 2)^3$ for both A and B.
- Algebraic multiplicity is 3.
- For A, three eigenvectors. Say, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$.
- For B, only one eigenvector ($\alpha\mathbf{e}_1$), so geometric multiplicity of B is 1.

Eigenvectors / Eigenvalues of Upper Triangular Matrix

Suppose A is upper triangular

$$A = \begin{bmatrix} A_{11} & \underline{u} & U_{13} \\ 0 & \lambda & \underline{v}^T \\ O & 0 & A_{33} \end{bmatrix}$$

Then

$$0 = (A - \lambda I)\underline{x} = \begin{bmatrix} U_{11} & \underline{u} & U_{13} \\ 0 & 0 & \underline{v}^T \\ O & 0 & U_{33} \end{bmatrix} \begin{pmatrix} \underline{y} \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} U_{11}\underline{y} - \underline{u} \\ 0 \\ 0 \end{pmatrix}$$

$$(A - \lambda I) \qquad \underline{x} \qquad\qquad 0$$

Because U_{11} is nonsingular, can solve $U_{11}\underline{y} = \underline{u}$ to find eigenvector $\underline{x}$.

Block Triangular Form

- If

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1p} \\ & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2p} \\ & & \ddots & \vdots \\ & & & \boldsymbol{A}_{pp} \end{bmatrix}$$

  with *square* diagonal blocks, then

$$\lambda(\boldsymbol{A}) = \bigcup_{j=1}^{p} \lambda(\boldsymbol{A}_{jj})$$

  so eigenvalue problem breaks into p smaller eigenvalue problems

- *Real* Schur form has 1×1 diagonal blocks corresponding to real eigenvalues and 2×2 diagonal blocks corresponding to pairs of complex conjugate eigenvalues

Eigenvalue-Revealing Factorizations

- Diagonalization: $A = X\Lambda X^{-1}$ if A is nondefective.
- Unitary diagonalization: $A = Q\Lambda Q^*$ if A is normal.
- Unitary triangularization: $A = QTQ^*$ always exists. (T upper triangular.)

Forms Attainable by Similarity

A	T	B
distinct eigenvalues	nonsingular	diagonal
real symmetric	orthogonal	real diagonal
complex Hermitian	unitary	real diagonal
normal	unitary	diagonal
arbitrary real	orthogonal	real block triangular (real Schur)
arbitrary	unitary	upper triangular (Schur)
arbitrary	nonsingular	almost diagonal (Jordan)

Always exists

- Given matrix A with indicated property, matrices B and T exist with indicated properties such that $B = T^{-1}AT$
- If B is diagonal or triangular, eigenvalues are its diagonal entries
- If B is diagonal, eigenvectors are columns of T

Similarity Transformations

- Given

$$B = T^{-1} A T$$
$$A = T B T^{-1}$$

- If A is normal ($AA^H = A^H A$),

$$A = Q \Lambda Q^H$$

  B is diagonal, T is unitary ($T^{-1} = T^H$).

- If A is symmetric real,

$$A = Q \Lambda Q^T$$

  B is diagonal, T is orthogonal ($T^{-1} = T^T$).

- If B is diagonal, T is the matrix of eigenvectors.

Computing Eigenpairs via Various (Sophisticated!) Forms of Power Iteration