Plan général

Motivations for eigenvalue / eigenvector computation

Mechanics and physics: Eigenvalues/eigenvectors useful for a very diverse array of reasons.

- Stability in time of mechanical and physical systems determined by eigenvalue problems. Likewise, eigenvalues reveal potential resonances.
- Eigenvectors often used (e.g. in structural dynamics) for low-dimensional approximation of dynamical responses.

Example: perturbation of dynamical system about equilibrium solution x_0 ($x(t)$, $x_0 \in \mathbb{R}^n$)

$$\dot{x} = \mathcal{F}(x), \qquad x(t) = x_0 + y(t),\ \mathcal{F}(x_0) = 0 \quad \Longrightarrow \quad \dot{y} = \mathcal{F}'(x_0)y + o(\|y\|)$$

Try $y(t) = Ye^{\lambda t}$ on linearized model, then $\lambda Y = \mathcal{F}'(x_0)Y$, stability if $\text{Re}(\lambda) < 0$.

Statistics: eigenvalues/eigenvectors of covariance matrices

Computation and algorithms:

- Computing matrix SVDs ($A = USV^{\text{H}}$) requires eigenvalues/eigenvectors. Then, singular values/vectors quantify information content of linear system.
- Knowing eigenvalue properties essential in *preconditioning* of iterative solution methods (lecture 4).

This lecture:

- A few major ideas and methods for eigenvalue/eigenvector computation;
- Focus on symmetric (Hermitian) eigenvalue problem;
- First, methods for isolated eigenvalues
- Then, methods for complete matrix spectra

Summary of facts about eigenvalues

For any $A \in \mathbb{K}^{n\times n}$, there exist n eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ and eigenvectors $x_1, \ldots, x_n$ such that

$$\boxed{Ax_i = \lambda_i x_i \qquad 1 \le i \le n.}$$

Eigenvalues: the n roots of the n-th degree *characteristic polynomial* $p_A(\lambda) := \det(A - \lambda I)$; can have multiplicities.

- If $\mathbb{K}^n = \text{span}(x_1, \ldots, x_n)$, A is *diagonalizable*: $A = X\Lambda X^{-1}$ ($\Lambda := \text{diag}(\lambda_1, \ldots, \lambda_n)$, $X := [x_1, \ldots, x_n]$).
- If $A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathsf{H}}$ with Q unitary, A is *unitarily diagonalizable*.
- Unitary diagonalization of A possible if and only if A *normal* ($AA^{\mathsf{H}} = A^{\mathsf{H}}A$).
- A Hermitian ($A = A^{\mathsf{H}}$) is normal, hence unitarily diagonalizable, and $\lambda_i \in \mathbb{R}$. A real symmetric $\implies$ Q orthogonal, $A = Q\Lambda Q^{\mathsf{T}}$.
- A not diagonalizable is called *defective*. Example:
$$A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}: \qquad \lambda_1 = \lambda_2 = a, \quad E_\lambda = \text{span}(x_1),\ x_1 = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$$
Algebraic multiplicity $\ge$ geometric multiplicity (here $2 \ge 1$).
- Gershgorin theorem: Each λ_k in one of the disks $\mathcal{D}_i(A) := \{ z \in \mathbb{C} : |z - a_{ii}| \le \sum_{j\neq i} |a_{ij}| \}$.

Convention: eigenvalues ordered such that $|\lambda_1| \ge |\lambda_2| \ge \ldots |\lambda_n|$.

Impossibility of direct algorithm for eigenvalue computation

- (a) Eigenvalues are roots of (degree-n) characteristic polynomial.
  (b) Any (degree-n) polynomial is characteristic polynomial of some matrix:

  $$P(X) = a_0 + a_1 X + \ldots + a_{n-1} X^{n-1} + X^n$$

  $$A_P = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}. \qquad \textit{companion matrix of } P$$

  Therefore eigenvalue problems and polynomial root-finding problems are **equivalent**
- No direct general root-finding method if $n \geq 5$ (impossibility result, Galois)
- Therefore **no direct method for general eigenvalue problems**

Any general method for computing matrix eigenvalues must be iterative.

This lecture:

- Computation of isolated eigenvalues
- Computation of matrix spectra

Computation of isolated eigenvalues: Rayleigh quotient

Computation of isolated eigenvalues: main ingredients are

- Power iterations
- Rayleigh quotients
- Matrix shifts

Let $A \in \mathbb{K}^{n\times n}$ Hermitian, $x \in \mathbb{K}^n$. Rayleigh quotient $r(x)$: $r(x) := \dfrac{x^{\mathsf{H}}Ax}{x^{\mathsf{H}}x}$.

We have $$\boxed{\nabla_x r(x) = \frac{2}{x^{\mathsf{H}}x}\big(Ax - r(x)x\big)}.$$

- Rayleigh quotient is stationary ($\nabla_x r(x) = 0$) if $x, r(x)$ eigenvector/eigenvalue pair.
- The smallest (largest) eigenvalue of A minimizes (maximizes) $r(x)$ over $x \in \mathbb{K}^n \setminus \{0\}$.
- Infinite-dimensional counterpart: Courant-Fischer min-max principle (ENSTA ANA 202)
- Dynamics of mechanical systems: $r(x)$ ratio of strain and kinetic energies.

Computation of isolated eigenvalues: power iterations

Repeated evaluations $x^{(0)} \mapsto Ax^{(0)} \mapsto A^2x^{(0)} \mapsto \ldots$, compute Rayleigh quotients along the way:

Algorithm 9 Power iteration

$A \in \mathbb{K}^{n\times n}$ Hermitian (input), $x^{(0)} \in \mathbb{K}^n$ with $\|x^{(0)}\| = 1$ (initialization)
for $k = 0, 1, 2, \ldots$ **do**
  $v = Ax^{(k)}$ (apply A to current normalized iterate)
  $\lambda^{(k)} = v^{\mathsf{H}}x^{(k)}$ (Rayleigh quotient)
  $x^{(k+1)} = v/\|v\|$ (next normalized iterate)
  Stop if convergence, set $\lambda_1 = \lambda^{(k)}$, $q_1 = x^{(k)}$
end for

Power iterations promote λ_1, q_1 (assuming $|\lambda_2| > |\lambda_1|$): use $A = Q\Lambda Q^{\mathsf{H}}$, expand on eigenvectors:

$$
\begin{aligned}
x^{(0)} &= y_1q_1 + y_2q_2 + \cdots + y_nq_n \\
x^{(k)} &= c_kA^kx^{(0)} \\
&= c_k\lambda_1^k\big[y_1q_1 + y_2(\lambda_2/\lambda_1)^kq_2 + \cdots + y_n(\lambda_n/\lambda_1)^kq_n\big]
\end{aligned}
$$

Convergence of power iterations

Assume $q_1^{\mathsf{H}}x^{(0)} \neq 0$ and $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$. Then:

$$|\lambda^{(k)} - \lambda_1| = O(|\lambda_2/\lambda_1|^{2k}), \qquad \|\pm x^{(k)} - q_1\| = O(|\lambda_2/\lambda_1|^k) \quad (k \to \infty)$$

Computation of isolated eigenvalues: power iterations (properties, limitations)

$$|\lambda^{(k)} - \lambda_1| = O(|\lambda_2/\lambda_1|^{2k}), \qquad \|\pm x^{(k)} - q_1\| = O(|\lambda_2/\lambda_1|^{k}) \quad (k \to \infty)$$

- Convergence of λ^k to λ_1 linear (error reduced by constant factor at each iteration);
- Convergence of $x^{(k)}$ to q_1 also linear.
- Error reduction depends on closeness of $|\lambda_1|$ and $|\lambda_2|$.
- "raw" power iteration, evaluate λ_1, q_1 only.
- Computational work: one matrix-vector product per power iteration.

Computation of isolated eigenvalues: inverse iterations

Idea: if A invertible, can apply power iterations to A^{-1}.

- Iterates: $x^{(k+1)} = A^{-1}x^{(k)}$ (i.e. solve $Ax^{(k+1)} = x^{(k)}$)
- Expected to produce λ_n, q_n (smallest eigenvalue)

Extension: apply power iterations to shifted matrix $(A-\mu I)^{-1}$, μ close to some λ_j:

- $\sigma_j := (\lambda_j - \mu)^{-1}$ an eigenvalue of $(A-\mu I)^{-1}$ with same eigenvector q_j; moreover
$$\boxed{|\mu - \lambda_j| < |\mu - \lambda_i| \ (i \neq j) \qquad \Longrightarrow \ |\sigma_j| > |\sigma_i| \ (i \neq j)}$$
- Power iterations on $(A-\mu I)^{-1} \implies (\sigma_j, q_j) \implies \lambda_j = \sigma_j^{-1} + \mu$
  Finds eigenvalue closest to μ

Algorithm 10 Inverse iteration

$A \in \mathbb{K}^{n\times n}$ Hermitian (input), $x^{(0)} \in \mathbb{K}^n$ with $\|x^{(0)}\| = 1$ (initialization), $\mu \in \mathbb{R}$ close to λ_j
for $k = 1, 2, \ldots$ **do**
  solve $(A-\mu I)x^{(k)} = x^{(k-1)}$ (apply $(A-\mu I)^{-1}$ to $x^{(k-1)}$)
  $x^{(k)} = x^{(k)}/\|x^{(k)}\|$ (next normalized iterate)
  $\lambda_j^{(k)} = (x^{(k)})^{\mathsf{H}}Ax^{(k)}$ (Rayleigh quotient)
  Stop if convergence, set $\lambda_j = \lambda_j^{(k)}$, $q = x^{(k)}$
end for

$$|\lambda_j^{(k)} - \lambda_j| = O\Big(\frac{|\lambda_j - \mu|}{|\lambda_\ell - \mu|}\Big)^{2k}, \qquad \|\pm x_k - q_j\| = O\Big(\frac{|\lambda_j - \mu|}{|\lambda_\ell - \mu|}\Big)^{k}$$

(similarly to power iterations)

Computation of isolated eigenvalues: Rayleigh quotient iterations

- Power iterations yield eigenvectors
- Inverse iterations yield eigenvalues
- Combine both: Rayleigh quotient iterations (set shift μ to current eigenvalue estimate)

Algorithm 11 Rayleigh quotient iteration

$A \in \mathbb{K}^{n\times n}$ Hermitian (input), $x^{(0)} \in \mathbb{K}^n$ with $\|x_0\| = 1$ (initialization)
$\lambda^{(0)} = (x^{(0)})^{\mathsf{H}} A x^{(0)}$ (Initialize Rayleign quotient)
for $k = 1, 2, \ldots$ **do**
  solve $(A - \lambda^{(k-1)} I) x^{(k)} = x^{(k-1)}$ (apply $(A - \lambda^{(k-1)} I)^{-1}$ to $x^{(k-1)}$)
  $x^{(k)} = x^{(k)} / \|x^{(k)}\|$ (next normalized iterate)
  $\lambda^{(k)} = (x^{(k)})^{\mathsf{H}} A x^{(k)}$ (Rayleigh quotient)
  Stop if convergence, set $\lambda = \lambda^{(k)}$, $q = x^{(k)}$
end for

Convergence of Rayleigh quotient iterations

Rayleigh iterations converge for almost all starting vectors $x^{(0)}$. Assume (normalized) $x^{(0)}$ close to eigenvector q_j. Then:

$$|\lambda^{(k+1)} - \lambda_j| = O(|\lambda^{(k)} - \lambda_j|^3), \qquad \| \pm x^{(k+1)} - q_j \| = O(\| \pm x^{(k)} - q_j \|^3) \quad (k \to \infty).$$

Computation of isolated eigenvalues: example

$A \in \mathbb{R}^{10\times 10}$ a random real symmetric matrix.

1. Power iterations for $\lambda_1 \approx 10.681$. Tolerance: $|\lambda^{(k+1)} - \lambda^{(k)}| \leq 10^{-10}$

k	$\lvert\lambda^{(k)} - \lambda_1\rvert$	$(\lambda_2/\lambda_1)^{2k}$	$\lVert q^{(k)} - q_1\rVert$	$(\lambda_2/\lambda_1)^k$
1	2.5383e+00	3.9667e-02	5.1770e-01	1.9916e-01
2	4.4221e-02	1.5734e-03	6.6508e-02	3.9667e-02
3	9.2950e-04	6.2413e-05	9.6536e-03	7.9002e-03
4	2.5329e-05	2.4757e-06	1.5871e-03	1.5734e-03
5	7.8583e-07	9.8202e-08	2.7731e-04	3.1337e-04
6	2.5948e-08	3.8953e-09	4.9888e-05	6.2413e-05
7	8.8738e-10	1.5451e-10	9.1258e-06	1.2430e-05
8	3.1084e-11	6.1290e-12	1.6892e-06	2.4757e-06
9	1.1084e-12	2.4312e-13	3.1567e-07	4.9307e-07

Computation of isolated eigenvalues: example

1. Inverse iterations ($(A-\mu I)^{-1}$ with $\mu = 0.4$ close to $\lambda_5 \approx 0.46740$).

k	$\lvert\lambda^{(k)} - \lambda_5\rvert$	$\left\lvert\frac{\lambda_5-\mu}{\lambda_4-\mu}\right\rvert^{2k}$	$\lVert q^{(k)} - q_5\rVert$	$\left\lvert\frac{\lambda_5-\mu}{\lambda_4-\mu}\right\rvert^{k}$
1	3.2395e-01	9.9940e-02	1.0111e+00	3.1613e-01
2	8.1309e-03	9.9880e-03	1.3554e-01	9.9940e-02
3	3.8992e-05	9.9819e-04	1.3551e-02	3.1594e-02
4	6.6151e-07	9.9759e-05	2.7233e-03	9.9880e-03
5	9.4566e-08	9.9699e-06	8.1688e-04	3.1575e-03
6	9.6321e-09	9.9639e-07	2.5726e-04	9.9819e-04
7	9.6379e-10	9.9579e-08	8.1308e-05	3.1556e-04
8	9.6329e-11	9.9519e-09	2.5704e-05	9.9759e-05
9	9.6276e-12	9.9459e-10	8.1257e-06	3.1537e-05

2. Rayleigh iterations for $\lambda_5 \approx 0.46740$.

k	$\lvert\lambda^{(k)} - \lambda_5\rvert$
1	6.6642e-03
2	7.0746e-07
3	4.9960e-16

Extension: generalized symmetric eigenvalue problems

- Many engineering applications (e.g. vibrations, forced dynamical motions) involve

  Find λ, x such that $\quad Kx = \lambda Mx.$

  $K, M \in \mathbb{R}^{n\times n}$: (symmetric, positive) stiffness and mass matrices.
- Usually, lower part of spectrum (for which FE model most accurate) sought
- Assume M SPD (true for dynamics/vibrations), set $M = GG^{\mathsf{T}}$ (Cholesky); then:

  $$Kx - \lambda Mx = G\big(A - \lambda I\big)G^{\mathsf{T}}x, \qquad A = G^{-1}KG^{-\mathsf{T}}$$
- Equivalent symmetric eigenvalue problem:

  Find λ, x such that $\quad Ay = \lambda y, \quad G^{\mathsf{T}}x = y.$

  Do not actually evaluate $A = G^{-1}KG^{-\mathsf{T}}$; reinterpret algorithms for standard eigenvalue pbs.

Algorithm 12 Inverse iteration for structural vibrations

$K, M \in \mathbb{R}^{n\times n}$ (stiffness / mass matrices), $x^{(0)} \in \mathbb{R}^n$, $\|x^{(0)}\| = 1$ (initialization), $\mu \in \mathbb{R}$ close to λ_j

for $k = 1, 2, \ldots$ **do**

solve $(K - \mu M)x^{(k)} = Mx^{(k-1)}$ (solve forced vibration problem with $Mx^{(k-1)}$ as load)

$x^{(k)} = x^{(k)}/\sqrt{x^{(k)\mathsf{T}}Mx^{(k)}}$ (next mass-normalized iterate)

$\lambda_j^{(k)} = (x^{(k)})^{\mathsf{H}}Kx^{(k)}$ (Rayleigh quotient)

Stop if convergence, set $\lambda_j = \lambda_j^{(k)}$, $q = x^{(k)}$

end for