Overview

- Matrix eigenvalue problem
- Reduction to upper Hessenberg form
- Power methods (Rayleigh quotient, inverse iteration)
- Example: cell cytoskeleton deformation modes

- $\boldsymbol{A} \in \mathbb{C}^{m \times m}$, $\mathbb{C}^m \xrightarrow{\boldsymbol{A}} \mathbb{C}^m$, $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$ (eigenvalue relationship), $\boldsymbol{x} \neq 0$.

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}, \boldsymbol{x} \neq \boldsymbol{0} \Rightarrow \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0 = p_{\boldsymbol{A}}(\lambda)$$

- $p_{\boldsymbol{A}}(\lambda)$ is of degree m, has m roots
  - n distinct roots, each appearing m_i times, (*algebraic multiplicity*) $\sum_{i=1}^{n} m_i = m$
  - $E_\lambda = \mathcal{N}(\boldsymbol{A} - \lambda \boldsymbol{I})$, $n_i = \dim E_{\lambda_i}$, $i = 1, \ldots, n$ (*geometric multiplicity*)
  - $E_i = \mathcal{N}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$, (*eigenspace* of λ_i)

Proposition. *Consider* $\boldsymbol{A} \in \mathbb{C}^{m \times m}$, $\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$, $\boldsymbol{A}\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \neq \boldsymbol{0}$. *If* $\lambda_1 \neq \lambda_2$ *then* $\boldsymbol{x}_1, \boldsymbol{x}_2$ *are linearly independent.*

Proof. *Consider* $c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2 = \boldsymbol{0}$, $c_1 \neq 0$, $\boldsymbol{x}_1 = -(c_2 / c_1)\boldsymbol{x}_2$. *From* $\boldsymbol{A}(c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2) = \boldsymbol{0}$

$$c_1 \lambda_1 \boldsymbol{x}_1 + c_2 \lambda_2 \boldsymbol{x}_2 = \boldsymbol{0} \Rightarrow c_2(\lambda_1 - \lambda_2)\boldsymbol{x}_2 = \boldsymbol{0} \Rightarrow c_2 = 0 \Rightarrow \boldsymbol{x}_1 = \boldsymbol{0} \text{ (contradiction)}$$

□

- Organize eigenvectors $\boldsymbol{X} = (\ \boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{x}_m\ ), \boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_m)$

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{\Lambda}$$

$$\boldsymbol{A}(\ \boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{x}_m\ ) = (\ \boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{x}_m\ )\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix}$$

- Let $\boldsymbol{E}_i \in \mathbb{C}^{m \times n_i}$ denote a basis of $E_i = \mathcal{N}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$. Recall that n_i is the geometric multiplicity of λ_i, i.e., $n_i = \dim(E_i)$.

$$\boldsymbol{A}(\ \boldsymbol{E}_1\ \ \ldots\ \ \boldsymbol{E}_n\ ) = (\ \boldsymbol{E}_1\ \ \ldots\ \ \boldsymbol{E}_n\ )\begin{pmatrix} \Lambda_1 & & \\ & \ddots & \\ & & \Lambda_n \end{pmatrix}, \Lambda_i = \begin{pmatrix} \lambda_i & & \\ & \ddots & \\ & & \lambda_i \end{pmatrix} \in \mathbb{C}^{m_i \times m_i}$$

- We may have $n_i < m_i$, λ_i is said to be defective
- $\boldsymbol{A}$ is non-defective or diagonalizable when $n_i = m_i$ for $i = 1, \ldots, n$. $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$, *eigendecomposition* of $\boldsymbol{A}$. Consider $\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1})\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}\boldsymbol{y}$.

- $\boldsymbol{Y} \in \mathbb{C}^{m \times m}$, non-singular, then the map from $\boldsymbol{A} \to \boldsymbol{Y}^{-1}\boldsymbol{A}\boldsymbol{Y}$ is a *similarity transform*.

$$p_{\boldsymbol{Y}^{-1}\boldsymbol{A}\boldsymbol{Y}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{Y}^{-1}\boldsymbol{A}\boldsymbol{Y}) = \det[\boldsymbol{Y}^{-1}(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{Y}] = p_{\boldsymbol{A}}(\lambda),$$

  $\boldsymbol{Y}^{-1}\boldsymbol{A}\boldsymbol{Y}$ has same eigenvalues as $\boldsymbol{A}$.

- $p_{\boldsymbol{A}}(\lambda) = (\lambda - \lambda_1)\ldots(\lambda - \lambda_m) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) \Rightarrow (-1)^m \lambda_1 \ldots \lambda_m = \det(-\boldsymbol{A}) = (-1)^m \det(\boldsymbol{A})$.
- $\boldsymbol{tr}(\boldsymbol{A}) = \sum_{i=1}^m a_{ii}$,

$$p_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda - a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & \lambda - a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & \lambda - a_{mm} \end{vmatrix} = \lambda^m - \mathrm{tr}(\boldsymbol{A})\lambda^{m-1} + \cdots +$$

$$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^m \lambda_i, \det(\boldsymbol{A}) = \prod_{i=1}^m \lambda_i.$$

- $\boldsymbol{Q} \in \mathbb{C}^{m \times m}$, unitary, $\boldsymbol{A} \to \boldsymbol{Q}^* \boldsymbol{A} \boldsymbol{Q}$ is a *unitary similarity transform*.

- $\boldsymbol{A}$ is *unitarily diagonalizable* if there exists some unitary matrix $\boldsymbol{Q}$ such that $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^*$.
- $\boldsymbol{A}$ is *normal* if $\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A}$

Theorem. *A matrix is unitarily diagonalizable iff it is normal.*

Theorem. *(Schur factorization) For any $\boldsymbol{A} \in \mathbb{C}^{m\times m}$, there exists $\boldsymbol{Q}$ unitary such that*

$$\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{T}, \boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^*$$

with $\boldsymbol{T}$ upper triangular.

- Summary:
  - $\boldsymbol{A}$ nondefective $\Rightarrow \boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$ (diagonalizable)
  - $\boldsymbol{A}$ normal $\Rightarrow \boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^*$ with $\boldsymbol{Q}\boldsymbol{Q}^* = \boldsymbol{I}$ (unitarily diagonalizable)
  - Arbitrary $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^*$, with $\boldsymbol{Q}$ unitary, $\boldsymbol{T}$ upper triangular.

- Given $\boldsymbol{A} \in \mathbb{C}^{m \times m}$, $\boldsymbol{Q}_{m-2}^* \ldots \boldsymbol{Q}_1^* \boldsymbol{A} \boldsymbol{Q}_1 \ldots \boldsymbol{Q}_{m-2} = \boldsymbol{H}$

Algorithm : Householder similarity reduction to Hessenberg form

Given $A \in \mathbb{R}^{m \times m}$
$H = A, V = 0_m$
for $j = 1 \colon m-2$
  $x = H(j+1 \colon m, j)$; $e = I_{m-j}(:, 1)$
  $H(j+1 \colon m, j) = \|x\| \cdot e$
  $v = \mathrm{sign}(x_1) \cdot H(j+1 \colon m, j) + x$; $v = v / \|v\|$
  $V(j+1 \colon m, j) = v$
  for $k = j+1 \colon m$
   $H(j+1 \colon m, k) = H(j+1 \colon m, k) - 2 \cdot v \cdot (v^* \cdot H(j+1 \colon m, k))$
   $H(k, j+1 \colon m) = H(k, j+1 \colon m) - 2 \cdot (H(k, j+1 \colon m) \cdot v) \cdot v^*$
  end
end
Return $[H, V]$

- $\mathcal{L} = \mathcal{K} + \mathcal{E} = \frac{1}{2}\dot{\boldsymbol{x}}^T \boldsymbol{M}\dot{\boldsymbol{x}} + \frac{1}{2}\boldsymbol{x}^T \boldsymbol{K}\boldsymbol{x}$, $\boldsymbol{M} = \boldsymbol{M}^T$, $\boldsymbol{K} = \boldsymbol{K}^T$
- Rayleigh quotient

$$r(\boldsymbol{x}) = \frac{\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}, \text{for}\, \boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}, r(\boldsymbol{x}) = \lambda.$$

- Taylor series

$$r(\boldsymbol{x}) = r(\boldsymbol{q}) + \nabla r(\boldsymbol{q}) \cdot (\boldsymbol{x} - \boldsymbol{q}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{q})^T \nabla \nabla r(\boldsymbol{q})(\boldsymbol{x} - \boldsymbol{q}) + \cdots$$

- Gradient of Rayleigh quotient

$$\nabla r(\boldsymbol{x}) = \frac{2}{\boldsymbol{x}^T \boldsymbol{x}}(\boldsymbol{A}\boldsymbol{x} - r(\boldsymbol{x})\boldsymbol{x}), \nabla r(\boldsymbol{q}) = 0$$

- Quadratically accurate

$$r(\boldsymbol{x}) - r(\boldsymbol{q}) = \mathcal{O}(\|\boldsymbol{x} - \boldsymbol{q}\|^2).$$

- $\boldsymbol{v}^{(0)} = a_1\boldsymbol{q}_1 + \cdots + a_m\boldsymbol{q}_m = \boldsymbol{Q}\,\boldsymbol{a}$, $\boldsymbol{A}\,\boldsymbol{q}_j = \lambda_j\,\boldsymbol{q}_j$

$$\begin{array}{l} \boldsymbol{v}^{(1)} = \boldsymbol{A}\,\boldsymbol{v}^{(0)} = a_1\lambda_1\boldsymbol{q}_1 + \cdots + a_m\lambda_m\boldsymbol{q}_m \\ \ldots \\ \boldsymbol{v}^{(k)} = \boldsymbol{A}\,\boldsymbol{v}^{(k-1)} = a_1\lambda_1^k\boldsymbol{q}_1 + \cdots + a_m\lambda_m^k\boldsymbol{q}_m \end{array} \quad \boldsymbol{v}^{(k)} = \lambda_1^k\left[a_1\boldsymbol{q}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k a_2\boldsymbol{q}_2 + \cdots + \left(\frac{\lambda_m}{\lambda_1}\right)^k a_m\boldsymbol{q}_m\right]$$

- After $k \gg 1$ iterations $\boldsymbol{v}^{(k)} \cong \lambda_1^k a_1\boldsymbol{q}_1$. $\boldsymbol{v}^{(k-1)} \cong \lambda_1^{k-1} a_1\,\boldsymbol{q}_1$. $\lambda_1 \cong v_i^{(k)} / v_i^{(k-1)}$, slower (linear convergence), less accurate (different estimates from each component). Better

$$\lambda_1^{(k)} \cong r(\boldsymbol{v}^{(k)} / \|\boldsymbol{v}^{(k)}\|), \text{quadratic convergence through Rayleigh quotient}$$

- Inverse iteration:
  - if λ is an eigenvalue of $\boldsymbol{A}$, then $\lambda - \mu$ is an eigenvalue of $\boldsymbol{A} - \mu\boldsymbol{I}$
  - if λ is an eigenvalue of $\boldsymbol{A}$, then λ^{-1} is an eigenvalue of $\boldsymbol{A}^{-1}$
  - if λ is an eigenvalue of $\boldsymbol{A}$, then $(\lambda - \mu)^{-1}$ is an eigenvalue of $(\boldsymbol{A} - \mu\boldsymbol{I})^{-1}$
  - apply power iteration to $(\boldsymbol{A} - \mu\boldsymbol{I})^{-1}$, $\boldsymbol{v}^{(k)} = (\boldsymbol{A} - \mu\boldsymbol{I})^{-1}\boldsymbol{v}^{(k-1)}$ implemented

$$(\boldsymbol{A} - \mu\boldsymbol{I})\boldsymbol{v}^{(k)} = \boldsymbol{v}^{(k-1)}.$$

- Combine inverse power iteration to find eigenvectors with Rayleigh quotient approximation of eigenvalues.

Algorithm Rayleigh quotient iteration (beautiful algorithm)

Given: $\boldsymbol{A} \in \mathbb{R}^{m \times m}$, $\boldsymbol{A}^T = \boldsymbol{A}$
$\boldsymbol{v}^{(0)}$, $\|\boldsymbol{v}^{(0)}\| = 1$, $\lambda^{(0)} = [\boldsymbol{v}^{(0)}]^T \boldsymbol{A} \boldsymbol{v}^{(0)}$
for $k = 1, 2, \ldots$
  Solve $(\boldsymbol{A} - \lambda^{(k-1)} \boldsymbol{I}) \boldsymbol{w} = \boldsymbol{v}^{(k-1)}$
  $\boldsymbol{v}^{(k)} = \boldsymbol{w} / \|\boldsymbol{w}\|$
  $\lambda^{(k)} = [\boldsymbol{v}^{(k)}]^T \boldsymbol{A} \boldsymbol{v}^{(k)}$

Theorem. *Rayleigh quotient iteration converges from almost all initial approximations* $(\lambda^{(0)}, \boldsymbol{v}^{(0)})$*. The asymptotic convergence rate is cubic*

$$|\lambda^{(k+1)} - \lambda| = \mathcal{O}(|\lambda^{(k)} - \lambda|^3)$$
$$\|\boldsymbol{v}^{(k+1)} - (\pm \boldsymbol{q})\| = \mathcal{O}(\|\boldsymbol{v}^{(k+1)} - (\pm \boldsymbol{q})\|^3)$$

$$W = -\int_0^u f(y)\,\mathrm{d}y = -\int_0^u ky\,\mathrm{d}y = -\tfrac{1}{2}ku^2$$
$$\mathcal{L} = K + W = \tfrac{1}{2}m\,\dot{u}^2 - \tfrac{1}{2}ku^2$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{u}}\right) - \left(\frac{\partial \mathcal{L}}{\partial u}\right) = m\ddot{u} + ku = 0.$$
$$\boldsymbol{r}_{ij}(t) = \boldsymbol{a}_j + \boldsymbol{u}_j(t) - \boldsymbol{a}_i + \boldsymbol{u}_i(t),\, \ell_{ij}(t) = \|\boldsymbol{r}_{ij}(t)\|$$
$$\boldsymbol{f}_{ij}(t) = k_{ij}(\ell_{ij}(t) - \ell_{ij}(0))\boldsymbol{r}_{ij}(t)\,/\,\ell_{ij}(t)$$
$$W_{ij} = \tfrac{1}{2}k_{ij}(\ell_{ij}(t) - \ell_{ij}(0))^2$$
$$W = \textstyle\sum_{i=1}^{n}\left(\sum'\right)_{j=i-p}^{i+p} W_{ij},\, W(\boldsymbol{u})$$
$$\boldsymbol{f} = \nabla_{\boldsymbol{u}} W(\boldsymbol{u}) \cong \boldsymbol{K}\boldsymbol{u}$$
$$\nabla_{\boldsymbol{u}} W(\boldsymbol{u}) = \nabla_{\boldsymbol{u}} W(\boldsymbol{0}) + \nabla_{\boldsymbol{u}}\nabla_{\boldsymbol{u}} W(\boldsymbol{0})\,\boldsymbol{u} + \ldots$$
$$\boldsymbol{K} = \nabla_{\boldsymbol{u}}\nabla_{\boldsymbol{u}} W(\boldsymbol{0})$$