Numerical Methods I: SVD and Orthogonal polynomials

Georg Stadler
Courant Institute, NYU
stadler@cims.nyu.edu

November 9, 2017
Summary of eigenvalue algorithms
Power method and inverse power method

▶ **Power method**: computes eigenvector for dominant eigenvalue (i.e., largest in absolute value)
▶ **Inverse power method**: computes eigenvector for eigenvalue closest to $\bar{\lambda}$

**Note**: Both methods only require repeated multiplication of the matrix $A$ (or its shifted inverse) to vectors, not the matrix $A$ itself.
Summary of eigenvalue algorithms

QR algorithm

- Computes all eigenvalues of $A$
- Each iteration requires a $QR$-factorization

To make it computationally efficient (for SPD matrices):

- Compute tridiagonal form using orthogonal transformation (Givens rotations or Householder)—$O(n^3)$ complexity
- In QR algorithm, tridiagonal matrices remain tridiagonal. Hence each step is $O(n^2)$.

**Note:** Requires explicit knowledge of $A$. Can be made efficient also for non-symmetric $A$ using upper Hessenberg forms.
Singular value decomposition

Let $A \in \mathbb{C}^{m \times n}$. The SVD decomposition is

$$A = U \Sigma V^*,$$

where $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$ are unitary/orthogonal, $\Sigma \in \mathbb{R}^{m \times n}$, $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p)$, $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_p \geq 0$, where $p = \min(m, n)$.

- If $A$ is real, it has a real SVD, i.e., $U, V$ can be chosen as real orthonormal matrices.
- The singular values are the square roots of the eigenvalues of $A^*A$.

$$A = U \Sigma V^* \Rightarrow AV = U \Sigma,$$

component-wise: $Av_i = \sigma_i u_i$,

$$V^*A^*AV = \Sigma^*U^*U\Sigma = \Sigma^*\Sigma = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_p^2 \end{pmatrix},$$

$$\Rightarrow A^*A V = V \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_p^2 \end{pmatrix} \Rightarrow \sigma_1^2 \ldots \sigma_p^2 \text{ EV of } A^*A \text{ eigenvalues are columns of } V.$$
Singular value decomposition

Existence

Induction: Show that there exist orthogonal/unitary matrices $U$, $V$ such that

$$U^T A V = \begin{bmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & B \end{bmatrix}$$

$$\alpha := \|A\|_2 = \max_{\|u\|_2 = 1} \|Au\|_2$$

$$Av = \alpha u, \quad v \in \mathbb{R}^n, \quad u \in \mathbb{R}^n$$

Extend $u, v$ to orthonormal bases

$$V = [v_1, v_2, \ldots, v_n] \in \mathbb{R}^{n \times n}$$

$$U = [u_1, u_2, \ldots, u_m] \in \mathbb{R}^{m \times m}$$

$$A_i = U^T A V = \begin{bmatrix} \alpha & U^T \omega \\ 0 & B \end{bmatrix}$$

Orthogonal/unitary

Remains to show $w = 0$

$$\|A_i (\omega)\|_2^2 \geq (\alpha^2 + \|w\|^2)^2 + \|Bw\|^2$$

$$\| (\omega)\|_2^2 = \alpha^2 + \|w\|^2$$
Singular value decomposition

Computation

\[ \alpha^2 = \| A \|_2^2 = \| A \|_2^2 \geq \frac{\| A_1 (\omega) \|_2^2}{\| (\omega) \|_2^2} \geq \frac{(\alpha^2 + \| w \|_2^2)^2}{\alpha^2 + \| w \|_2^2} = \alpha^2 + \| w \|_2^2 \]

\[ \rightarrow w = 0 \]

Computation: Assume \( m \geq n \)

 Singular values are square roots of eigenvalues of \( A^* A \) [all these eigenvalues are real and non-negative because \( A^* A \) is symmetric, positive semi-definite]
Singular value decomposition

Computation

First find $P,Q$ orthogonal such that $PAQ = \begin{bmatrix} \ast & \ast \\ 0 & \ast \end{bmatrix} = \begin{bmatrix} B \\ 0 \end{bmatrix}$

This can be done e.g. using Givens rotations:

\[
\begin{align*}
\begin{bmatrix} x \\ \vdots \\ x \\ \vdots \end{bmatrix} & \xrightarrow{P_i} \begin{bmatrix} x \\ \vdots \\ 0 \\ \vdots \end{bmatrix} \cdot Q_i \xrightarrow{\cdots} \begin{bmatrix} \ast & \ast & \ast & \ast \\ 0 & \ast & \ast & \ast \\ 0 & 0 & \ast & \ast \\ 0 & 0 & 0 & \ast \end{bmatrix} \\
\end{align*}
\]

eigenvalues of $A^*A$ are identical to eigenvalues of $B^*B$

because $B^*B = Q^*A^*P^*PAQ = Q^*A^*AQ$

$B^*B = \begin{bmatrix} \ast & \ast \\ 0 & 0 \end{bmatrix}$

tridiagonal matrix. We could compute $B^*B$

and then use QR algorithm to find its eigenvalues. However, we can perform the QR iterations for $B^*B$ only working with $B$, which avoids computation of $B^*B$. 


Overview

Orthogonal polynomials and 3-term recurrence relations
Inner products between vectors and functions

Let \( u_1, v_1, u_2, v_2 \in V, \alpha \in \mathbb{R} \)

inner product satisfies:

\( \langle u_1, u_2 \rangle \geq 0, \quad \langle u_1, u_2 \rangle = 0 \iff u = 0 \)

\( \langle u_1, v_1 \rangle = \langle v_1, u_1 \rangle \)

\( \langle \alpha u_1 + u_2, v \rangle = \alpha \langle u_1, v \rangle + \langle u_2, v \rangle \)

\( \| u \| := \sqrt{\langle u, u \rangle} \) induced norm

Examples in \( \mathbb{R}^n \):

\( \langle u, v \rangle = u^T v \) Euclidean inner product

induces 2-norm

\( \langle u, v \rangle_w := u^T W v, \) \( W \) is spd

weighted inner products
Orthogonal polynomials

Choices for inner products between functions:

$$(f, g) = \int_{-\pi}^{\pi} f(x)g(x) \, dx.$$  

Orthogonal functions w.r. to this inner product are $\cos(kx)$, $\sin(kx)$.

They satisfy a three-term recurrence relation:

$$T_k(x) = 2 \cos(x)T_{k-1}(x) - T_{k-2}(x),$$

where $T_k(x) = \cos(kx)$ or $T_k(x) = \sin(kx)$. 

Define an inner products between functions:

\[(f, g) = \int_a^b \omega(x)f(x)g(x) \, dx,\]

where \(\omega(x) > 0\) for \(a < x < b\) is a weight function. The induced norm is \(\|f\| := \sqrt{(f, f)}\).

Examples for \(\omega(x)\):

\(\omega(x) \equiv 1,\)

or

\(\omega(x) = \frac{1}{\sqrt{1 - x^2}}\) for \(a = -1, b = 1\).
Orthogonal polynomials

Denote by $P_k$ the polynomials of degree $k$.

**Theorem:** There exist uniquely determined orthogonal polynomials $P_k \in P_k$ with leading coefficient 1. These polynomials satisfy the 3-term recurrence relation:

$$P_k(t) = (t + a_k)P_{k-1}(t) + b_kP_{k-2}(t), \quad k = 2, 3$$

with starting values $P_0 = 1$, $P_1 = t + a_1$, where

$$a_k = -\frac{(tP_{k-1}, P_{k-1})}{(P_{k-1}, P_{k-1})}, \quad b_k = -\frac{(P_{k-1}, P_{k-1})}{(P_{k-2}, P_{k-2})}$$
Orthogonal polynomials

Sketch of proof:

\[ P_0 = 1, \quad P_1 = t, \quad P_0, \ldots, P_{k-1} \text{ constructed to be orthogonal} \]

\[ P_k(t) = t^k + \cdots + t^{k-1} + \ldots \]

\[ P_k - tP_{k-1} \text{ degree } \leq k-1 \quad \Rightarrow \quad P_k - tP_{k-1} = \sum_{j=0}^{k-1} c_j \cdot P_j. \]

\[ P_k \text{ must be orthogonal to } P_0, \ldots, P_{k-1} \]

\[ (P_k - tP_{k-1}, P_e) = \left( \sum_{j=0}^{k-1} c_j \cdot P_j, P_e \right) = c_l (P_{k-1}, P_e) \]

\[ l = 0, \ldots, k-3: \quad c_l = \frac{(P_k - tP_{k-1}, P_e)}{(P_{k-1}, P_e)} = 0 \quad \text{because } (P_k, P_e) = 0 \]

\[ l = k-2: \quad c_{k-2} = \frac{(P_{k-1} - tP_{k-2})}{(P_{k-1}, P_{k-2})} = 0 \quad \text{degree } \leq k-2 \]

\[ l = k-1: \quad c_{k-1} = \frac{(P_k - tP_{k-1}, P_{k-1})}{(P_{k-1}, P_{k-1})} = a_k \]

\[ P_k = (t + c_{k-1}) \cdot P_{k-1} + c_{k-2} \cdot P_{k-2} \]
Theorem: The orthogonal polynomials $P_k \in \mathbb{P}_k$ have exactly $k$ simple roots in $(a, b)$.

Note that 3-term recurrence can be used to
- Compute polynomials $P_k$ completely, or
- Evaluate $P_k$ at a point $x_0$. 
Orthogonal polynomials

Chebyshev polynomials

Chebyshev polynomials, for $-1 \leq x \leq 1$:

$$T_k(x) = \cos(k \arccos(x))$$

3-term recursion: $T_0(x) = 1$, $T_1(x) = x$,

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x) \quad \text{for } k \geq 2.$$  

$$T_2(x) = 2x^2 - 1$$

- $T_k$ are polynomials,
- $T_k$ can be defined for all $x \in \mathbb{R}$. 

Orthogonal polynomials

Chebyshev polynomials

In what inner product are Chebyshev polynomials orthogonal?

$$\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) \, dx = \begin{cases} 
0 & n \neq m \\
\pi & n = m = 0 \\
\pi/2 & n = m \neq 0
\end{cases}$$

The roots of Chebyshev polynomials play an important role in interpolation.
Legendre polynomials

Orthogonal polynomials with for weight function $\omega \equiv 1$, satisfy $L_0 = 1$, $L_1 = x$, and

$$L_{k+1}(x) = \frac{2k + 1}{k + 1} x L_k(x) - \frac{k}{k + 1} L_{k-1}(x)$$

Orthogonal polynomials on the sphere (defined in spherical coordinates $(\theta, \varphi)$; also satisfy a 3-term recursion

Numerical aspects

- Simple application of 3-term recurrence might not always be stable due to cancellation.
- Adjoint summation (Sec 6.3 in Deuflhard/Hohmann) can avoid cancellation errors.