# 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

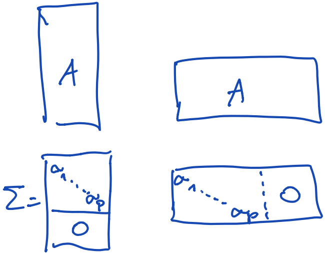
- Computes all eigenvalues of $A$
- Each iteration requires a $QR$-factorization

To make it computationally efficient (for SPD matrices):

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

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

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

$U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ 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 $A^*A$.

$A = U\Sigma V^* \Longrightarrow AV = U\Sigma$, componentwise: $Av_i = \sigma_i u_i$, $i = 1, \ldots p$

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

$\Longrightarrow A^*AV = V\begin{pmatrix} \sigma_1^2 \ldots \sigma_p^2 \\ 0 \end{pmatrix} \Longrightarrow \sigma_1^2, \ldots \sigma_p^2$ EV of $A^*A$, eigenvectors are columns of $V$

# Singular value decomposition

Existence

Induction: Show that $\exists\, U, V$ orthogonal / unitary such that

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

$$\sigma := \|A\|_2 = \max_{\|v\|_2 = 1} \|Av\|_2 \qquad Av = \sigma u, \quad v \in \mathbb{R}^n, u \in \mathbb{R}^m, \ \|v\| = \|u\| = 1$$

Extend $u, v$ to orthonormal bases

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

orthogonal / unitary

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

$$A_1 = U^T A V = \begin{bmatrix} \sigma & w^T \\ 0 & \\ \vdots & B \\ 0 & \end{bmatrix}$$

Remains to show $w = 0$

$$\left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2^2 \geq (\sigma^2 + \|w\|^2)^2 \cancel{+ \|Bw\|^2}$$

$$\left\| \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|^2 = \sigma^2 + \|w\|^2$$

# Singular value decomposition

## Computation

$$\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2 \geq \frac{\left\|A_1\binom{\sigma}{w}\right\|_2^2}{\left\|\binom{\sigma}{w}\right\|_2^2} \geq \frac{(\sigma^2+\|w\|^2)^2}{\sigma^2+\|w\|^2} = \sigma^2+\|w\|^2$$

$$\implies w = 0$$

---

Computation: Assume $m \geq n$ $\qquad A =$

Singular values are square roots of eigenvalues of $A^*A$ [all these eigenvalues are real and non-negative because $A^*A$ is symmetric, positive semidefinite]

# Singular value decomposition

## Computation

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

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

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

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

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

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

Orthogonal polynomials and 3-term recurrence relations

# Inner products between vectors and functions

Let $u, v, u_1, u_2 \in V$, $\alpha \in \mathbb{R}$

inner products satisfy:

- $(u,u) \geq 0$, $(u,u) = 0 \Leftrightarrow u = 0$
- $(u,v) = (v,u)$
- $(\alpha u_1 + u_2, v) = \alpha (u_1, v) + (u_2, v)$

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

Examples in $\mathbb{R}^n$: $(u,v) = u^T v$ Euclidean inner product, induces 2-norm

$(u,v)_W := u^T W v$, $W$ is spd
weighted inner product

# Orthogonal polynomials

Choices for inner products between functions:

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

Orthogonal functions w.r.t. 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 $T_k(x) = \sin(kx)$.

# Orthogonal polynomials

Define an inner products between functions:

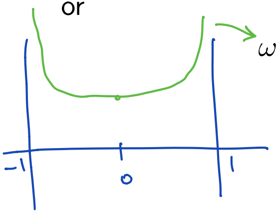
$$(f, g) = \int_a^b \omega(x) f(x) g(x)\, dx,$$

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




$$\omega(x) = \frac{1}{\sqrt{1 - x^2}} \quad \text{for} \quad a = -1, b = 1.$$

# Orthogonal polynomials

Denote by $\boldsymbol{P}_k$ the polynomials of degree $k$.

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

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

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: $P_0 \equiv 1$, $P_1 = t$, $P_0, \dots, P_{k-1}$ constructed to be orthogonal

$$P_k(t) = t^k + * \, t^{k-1} + \dots$$

$P_k - tP_{k-1}$ degree $\le k-1$ $\Longrightarrow$ $P_k - tP_{k-1} = \sum_{j=0}^{k-1} c_j \cdot P_j$.

$P_k$ must be orthogonal to $P_0, \dots, P_{k-1}$

$$\left(P_k - tP_{k-1}, P_\ell\right) = \left(\sum_{j=0}^{k-1} c_j \cdot P_j, P_\ell\right) = c_\ell \left(P_\ell, P_\ell\right)$$

$\ell = 0, \dots, k-3$: $c_\ell = \dfrac{(P_k - tP_{k-1}, P_\ell)}{(P_\ell, P_\ell)} = 0$ because $(P_k, P_\ell) = 0$, $(-tP_{k-1}, P_\ell) = (-P_{k-1}, \underbrace{tP_\ell}_{\text{degree} \le k-2}) = 0$

$\ell = k-2$: $c_{k-2} = \dfrac{(\cancel{P_k} - tP_{k-1}, P_{k-2})}{(P_{k-2}, P_{k-2})}$

$\ell = k-1$: $c_{k-1} = \dfrac{(\cancel{P_k} - tP_{k-1}, P_{k-1})}{(P_{k-1}, P_{k-1})}$ $\longrightarrow$ $P_k = (t + \overset{a_k}{c_{k-1}}) P_{k-1} + \overset{b_k}{c_{k-2}} P_{k-2}$

# Orthogonal polynomials

Theorem: The orthogonal polynomials $P_k \in \boldsymbol{P}_k$ 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 $-1 \leq x \leq 1$:

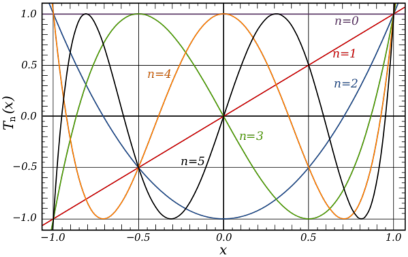
$$T_k(x) = \cos(k \arccos(x))$$

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 $x \in \mathbb{R}$.




Source: Wikipedia.

In what inner product are Chebyshev polynomials orthogonal?

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

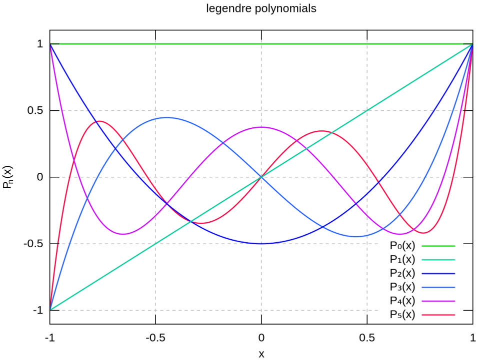
The roots of Chebyshev polynomials play an important role in interpolation.

# Orthogonal polynomials

## Legendre polynomials

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

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

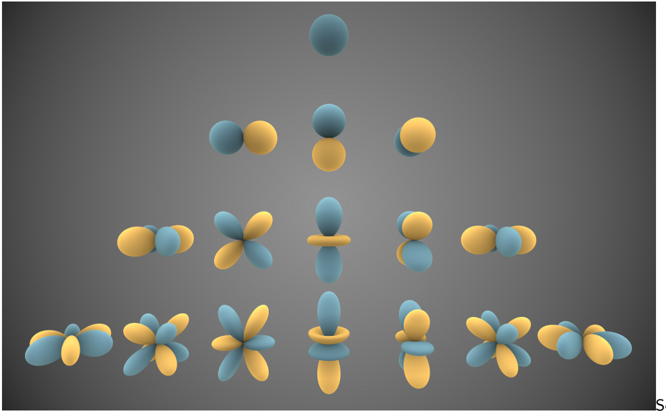



Source: Wikipedia.

# Orthogonal polynomials

## Spherical harmonics

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



Source: Wikipedia.

# 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.