September 9, 2014

Chapter 4: Linear Algebra Background

Uri M. Ascher and Chen Greif
Department of Computer Science
The University of British Columbia
{ascher,greif}@cs.ubc.ca

Slides for the book
A First Course in Numerical Methods (published by SIAM, 2011)
`http://bookstore.siam.org/cs07/`

Goals of this chapter

- To provide common background (no numerical algorithms) in linear algebra, necessary for developing numerical algorithms elsewhere;
- to collect several concepts and definitions for easy referencing;
- to ensure that those who have the necessary background can easily skip this chapter.

Outline

- Basic concepts: linear systems and eigenvalue problems
- Vector and matrix norms
- Symmetric positive definite and orthogonal matrices
- Singular value decomposition

Basic concepts: linear system of equations

- Find $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ which satisfies

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}$$

  or $A\mathbf{x} = \mathbf{b}$ with $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.
- Unique solution **iff** lines are not parallel.
- In general, for a square $n \times n$ system there is a unique solution if one of the following equivalent statements hold:
  - A is nonsingular;
  - $\det(A) \neq 0$;
  - A has linearly independent columns or rows;
  - there exists an inverse A^{-1} satisfying $AA^{-1} = I = A^{-1}A$;
  - $\mathrm{range}(A) = \mathbb{R}^n$;
  - $\mathrm{null}(A) = \{0\}$.

Basic concepts: eigenvalue problems

- A scalar λ and a vector $\mathbf{x}$ are an eigenvalue-eigenvector pair (or eigenpair) if

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

- For a *diagonalizable* $n \times n$ real matrix A there are n (generally complex-valued) eigenpairs $(\lambda_j, \mathbf{x}_j)$, with $X = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ nonsingular, and $X^{-1}AX$ is a diagonal matrix with the eigenvalues on the main diagonal.
- **Similarity transformation:** Given a nonsingular matrix S, the matrix $S^{-1}AS$ has the same eigenvalues as A. (Exercise: what about the eigenvectors?)

Outline

- Basic concepts: linear systems and eigenvalue problems
- Vector and matrix norms
- Symmetric positive definite and orthogonal matrices
- Singular value decomposition

Vector norms

A **vector norm** is a function "$\|\cdot\|$" from $\mathbb{R}^n$ to $\mathbb{R}$ that satisfies:

1. $\|\mathbf{x}\| \geq 0; \quad \|\mathbf{x}\| = 0 \text{ iff } \mathbf{x} = \mathbf{0}$,
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\| \;\; \forall\, \alpha \in \mathbb{R}$,
3. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

This generalizes absolute value or magnitude of a scalar.

Famous vector norms

- ℓ_2-*norm*

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.$$

- ℓ_∞-*norm*

$$\|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i|.$$

- ℓ_1-*norm*

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|.$$

Example

- Problem: Find the distance between

$$\mathbf{x} = \begin{pmatrix} 11 \\ 12 \\ 13 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} 12 \\ 14 \\ 16 \end{pmatrix}.$$

- Solution: let

$$\mathbf{z} = \mathbf{y} - \mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

and find $\|\mathbf{z}\|$.

- Calculate

$$\begin{aligned} \|\mathbf{z}\|_1 &= 1 + 2 + 3 = 6, \\ \|\mathbf{z}\|_2 &= \sqrt{1 + 4 + 9} \approx 3.7417, \\ \|\mathbf{z}\|_\infty &= 3. \end{aligned}$$

Matrix norms

Induced matrix norm of $m \times n$ matrix A for a given vector norm:

$$\|A\| = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \quad = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|.$$

Then consistency properties hold,

$$\|AB\| \le \|A\|\|B\|, \quad \|A\mathbf{x}\| \le \|A\|\|\mathbf{x}\|,$$

in addition to the previously stated three norm properties.

Famous matrix norms

- ℓ_2-*norm*

$$\|A\|_2 = \sqrt{\rho(A^T A)},$$

where ρ is **spectral radius**

$$\rho(B) = \max\{|\lambda|; \;\; \lambda \text{ is an eigenvalue of } B\}.$$

- ℓ_∞-*norm*

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|.$$

- ℓ_1-*norm*

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Outline

- Basic concepts: linear systems and eigenvalue problems
- Vector and matrix norms
- Symmetric positive definite and orthogonal matrices
- Singular value decomposition

Symmetric positive definite matrices

Extend notion of positive scalar to matrices:

$$A = A^T, \quad \mathbf{x}^T A \mathbf{x} > 0, \quad \text{all } \mathbf{x} \neq \mathbf{0}.$$

A symmetric matrix is positive definite if and only if all its eigenvalues are positive:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0.$$

Orthogonal matrices

Orthogonal vectors

Two vectors $\mathbf{u}$ and $\mathbf{v}$ of the same length are orthogonal if

$$\mathbf{u}^T\mathbf{v} = 0.$$

Orthonormal vectors: if *also* $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$.

Square matrix Q is **orthogonal** if its columns are pairwise orthonormal, i.e.,

$$Q^TQ = I. \quad \text{Hence also } Q^{-1} = Q^T.$$

Important property: for any orthogonal matrix Q and vector $\mathbf{x}$

$$\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2.$$

Hence

$$\|Q\|_2 = \|Q^{-1}\|_2 = 1.$$

Outline

- Basic concepts: linear systems and eigenvalue problems
- Vector and matrix norms
- Symmetric positive definite and orthogonal matrices
- Singular value decomposition

Singular value decomposition

Let A be real $m \times n$ (rectangular in general). Then there are orthogonal matrices $U,\ V$ such that

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

where

$$\Sigma = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}, \quad S = \mathrm{diag}\{\sigma_1, \ldots, \sigma_r\},$$

with the **singular values** $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, $\sigma_{r+1} = \cdots = \sigma_n = 0$.

Connection to eigenvalues: $\sigma_i = \sqrt{\lambda_i}$, where λ_i are eigenvalues of $A^T A$.

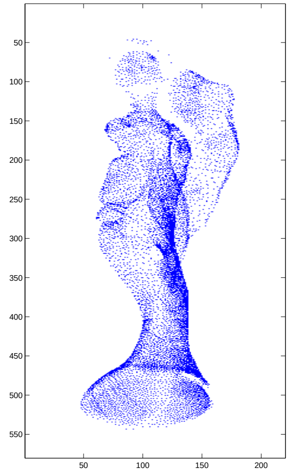
Example: principal component analysis (PCA)

Given a data matrix A each column corresponds to a different experiment of the same type in dimension m. Assume A has zero empirical mean.
PCA is an SVD transformation, rotating coordinates to align the transformed axes with the directions of maximum variance.
So $B = U^T A = \Sigma V^T$ is better than A. Covariance matrix

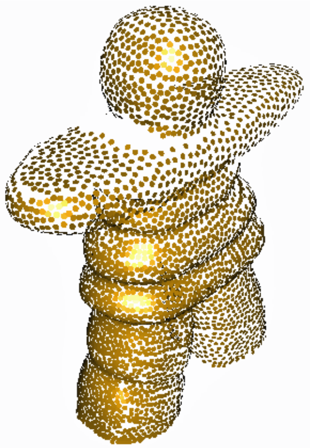
$$C = AA^T = U\Sigma\Sigma^T U^T.$$

Instance of use: **dimensionality reduction**. Let U_r consist of the first r columns of U, $r < n$.
Represent the data by the smaller matrix $B_r = U_r^T A$. Then $B_r = \Sigma_r V_r^T$.

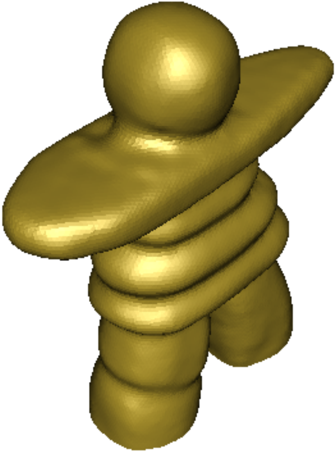
Instance: point cloud



Instance: RBF interpolation



(a) consolidated point cloud



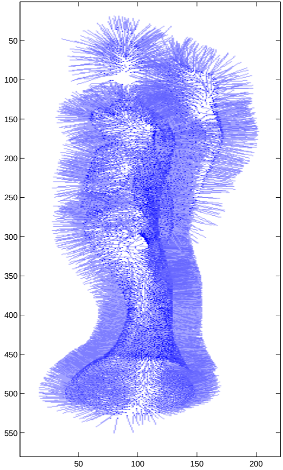
(b) RBF surface

FIGURE : RBF interpolation of an upsampling of a consolidated point cloud.

Normals to cloud points

- For a fixed $\mathbf{p}$ in the cloud, define a neighborhood $\mathcal{N}_p$ of nearby points.
- Calculate the mean of neighbors to find the centroid $\bar{\mathbf{p}}$.
- Then the $3 \times n_p$ data matrix A has $\mathbf{p}_{i_p} - \bar{\mathbf{p}}$ for its ith column.
- Find the three singular vectors of A (i.e. the eigenvectors of the covariance matrix C).
- The first two principal vectors span the **tangent plane** at $\mathbf{p}$. The third is the unsigned **normal direction**.

Point cloud with normals



Example: data fitting

Given measurements, or observations

$$(t_1, b_1), (t_2, b_2), \ldots, (t_m, b_m) = \{(t_i, b_i)\}_{i=1}^{m},$$

want to fit a function

$$v(t) = \sum_{j=1}^{n} x_j \phi_j(t),$$

- $\phi_1(t), \phi_2(t), \ldots, \phi_n(t)$ are known linearly independent **basis functions**
- $x_1, \ldots, x_n$ are **coefficients** to be determined s.t.

$$v(t_i) = b_i, \quad i = 1, 2, \ldots, m.$$

Data fitting cont.

Define $a_{ij} = \phi_j(t_i)$. Want $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Assume that A has full column rank n.

1. If $m = n$ get **interpolation problem**.
2. If $m > n$ want, e.g., $\min_{\mathbf{x}} \|\mathbf{b} - A\mathbf{x}\|_2$. Get **least squares data fitting**.

Example: differential equation

Given $g(t), \ 0 \le t \le 1$, recover $v(t)$ satisfying $-v'' = g$.
Require two boundary conditions

1. $v(0) = v(1) = 0$, or
2. $v(0) = 0,\ v'(1) = 0$.

Discretize on mesh $t_i = ih, \ i = 0, 1, \ldots, N$:

$$-\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} = g(t_i), \quad i = 1, 2, \ldots, N-1.$$

With BC $v(0) = v(1) = 0$, require $v_0 = v_N = 0$.

Linear system for differential equation

Need to solve $A\mathbf{v} = \mathbf{g}$, where

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{N-2} \\ v_{N-1} \end{pmatrix}, \ \mathbf{g} = \begin{pmatrix} g(t_1) \\ g(t_2) \\ \vdots \\ g(t_{N-2}) \\ g(t_{N-1}) \end{pmatrix}, \ A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}.$$

Thus, A is **tridiagonal**.