CSci 6974 and ECSE 6966 Math. Tech. for Vision, Graphics and Robotics
Lecture 7, February 9, 2006
SVD and Other Decompositions

Announcement

Question 2 of the "Problems for Grading" from Lecture 6 is canceled.

A Little More on Eigenvalues and Diagonalizations

- As discussed in Lecture 6, when a square matrix $\mathbf{A}$ has n distinct eigenvalues, $\lambda_1, \ldots, \lambda_n$ and associated eigenvectors, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then the eigenvectors are linearly independent and $\mathbf{A}$ may be written

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \tag{1}$$

  where

$$\mathbf{P} = \begin{pmatrix}\mathbf{v}_1, \ldots, \mathbf{v}_n\end{pmatrix} \qquad \text{and} \qquad \mathbf{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n). \tag{2}$$

  This is a *diagonalization*.

- Unfortunately, when the eigenvalues of $\mathbf{A}$ are non distinct, this diagonalization may not be possible. For example,

$$\mathbf{A} = \begin{pmatrix} 6 & -1 \\ 1 & 4 \end{pmatrix} \tag{3}$$

  has characteristic equation

$$\lambda^2 - 10\lambda + 25 = 0 \tag{4}$$

  which produces a double root at $\lambda = 5$. The associated eigenvector is $(1, 1)^\top$ and its multiples. This is the only eigenvector of $\mathbf{A}$.

  - The best we can do for this particular matrix is to write it in *Jordan Canonical Form* — a topic that is outside the scope of our discussion.

- When $\mathbf{A}$ is symmetric, its eigenvalues are real, and we can write

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^T, \tag{5}$$

  where $\mathbf{V}$ is an orthonormal matrix of eigenvectors and $\mathbf{D}$ is a diagonal matrix of eigenvalues. Moreover, if the eigenvalues are unique, this is unique up to a re-ordering of the columns of $\mathbf{V}$ and the associated entries of $\mathbf{D}$.

Motivation for Today's Class

We've already discussed the LU decomposition and the QR decomposition. Now we'll consider the *Cholesky* factorization and (especially) the *Singular Value Decomposition (SVD).*

Cholesky Factorization

- Any symmetric, positive definite matrix $\mathbf{A}$ may be factored into the product of an lower triangular matrix and its transpose:

  $$\mathbf{A} = \mathbf{L}\mathbf{L}^T$$

  and into the product of an upper triangular matrix and its transpose:

  $$\mathbf{A} = \mathbf{K}\mathbf{K}^T$$

  If we add the requirement that the diagonal entries are positive, then each factorization is unique.

- This follows from a combination of the spectral decomposition and the QR factorization. It also follows from the LU-decomposition, although we will not derive this in class.

SVD — Motivation

- Square matrices with distinct eigenvalues may be diagonalized (decomposed) using a matrix, $\mathbf{M}$, whose columns are the eigenvectors.
- Symmetric matrices may be decomposed using the spectral decomposition.
- What about arbitrary matrices? This is where the singular value decomposition (SVD) appears so spectacularly.

SVD — Outline of Derivation

(This derivation is adapted from Appendix A of Strang's, *Linear Algebra and Its Applications*, 3rd edition.)

- Let $\mathbf{A}$ be an $m \times n$ matrix and without losing generality assume $m \geq n$.
- Form $n \times n$ matrix $\mathbf{A}^\top\mathbf{A}$, and compute its eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. We'll assume that these are ordered so that $\lambda_i \geq \lambda_j$ when $i < j$.
- Observe that the eigenvalues λ_i are non-negative, and then define $\sigma_i = \sqrt{\lambda_i}$.
- Form the matrix

  $$\mathbf{V} = \big(\mathbf{v}_1, \ldots, \mathbf{v}_n\big). \tag{6}$$

- Define
$$\mathbf{u}_i = \frac{\mathbf{A}\mathbf{x}_i}{\sigma_i}. \tag{7}$$
- Several notes about the $\mathbf{u}_i$ vectors are important
  - The vectors $\mathbf{u}_i$ are m-dimensional, whereas the vectors $\mathbf{v}_i$ are n-dimensional.
  - The vectors in the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ are orthonormal.
- Extend the set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ to an orthonormal basis for $\mathbb{R}^m$ by finding $\{\mathbf{u}_{n+1}, \ldots, \mathbf{u}_m\}$ using Gram-Schmidt. (We can test the canonical basis to determine vectors to add.)
- Form these vectors into the matrix
$$\mathbf{U} = \big(\mathbf{u}_1, \ldots, \mathbf{u}_m\big). \tag{8}$$
  Note that $\mathbf{U}$ is orthogonal and $m \times m$, just as $\mathbf{V}$ is $n \times n$ and orthogonal.
- Form a $m \times n$ matrix $\mathbf{D}$ by placing the values σ_i in location $D_{i,i}$ and 0's elsewhere.
- Putting all of this together we can easily show that
$$\mathbf{D} = \mathbf{U}^\top \mathbf{A}\mathbf{V}. \tag{9}$$
- Therefore, we have
$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top. \tag{10}$$
  This the *Singular Value Decomposition* of $\mathbf{A}$.

SVD — 2nd Form

A more common and computational useful form of the SVD drops the last $m-n$ columns of $\mathbf{U}$ and the last $m-n$ rows of $\mathbf{D}$ to obtain the following form of the SVD.

Any real $m \times n$ matrix $\mathbf{A}$ may be decomposed as
$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{11}$$
where $\mathbf{U}$ is $m \times n$ and contains orthonormal columns, $\mathbf{V}$ is $n \times n$ and orthogonal, and $\mathbf{D}$ is a diagonal matrix of non-negative real numbers,
$$\mathbf{D} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \tag{12}$$
ordered so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. This decomposition is unique if the values of σ_i are distinct.

Properties

- Using the second version (11), and letting $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ and $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, then

$$\mathbf{A} = \sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T. \tag{13}$$

- Multiplying:

$$\begin{aligned} \mathbf{A}^T\mathbf{A} &= \mathbf{V}\mathbf{D}\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T \\ &= \mathbf{V}\mathbf{D}^2\mathbf{V}^T \\ &= \mathbf{V}\mathbf{D}^2\mathbf{V}^{-1} \end{aligned}$$

  This reverses our derivation of the SVD to show that the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ are the squares of the singular values of $\mathbf{A}$.

- Only $\mathbf{V}$ and $\mathbf{D}$ affect $\|\mathbf{A}\mathbf{x}\|$:

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|^2 &= \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} \\ &= \mathbf{x}^T\mathbf{V}\mathbf{D}^2\mathbf{V}^T\mathbf{x} \\ &= \|\mathbf{D}\mathbf{V}^T\mathbf{x}\|^2 \end{aligned}$$

- The SVD immediately gives us bases for the row-space, column-space, and nullspaces of $\mathbf{A}$. Assuming there are k non-zero singular values:
  - $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ form an orthonormal basis for the column space.
  - $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ form an orthonormal basis for the row space.
  - $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ form an orthonormal basis for the null space.
  - $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_m\}$ form an orthonormal basis for the left nullspace (using the first form of the SVD).

- For square, non-singular matrices:

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{D}^{-1}\mathbf{U}^T \tag{14}$$

  and, of course,

$$\mathbf{D}^{-1} = \mathrm{diag}(1/\sigma_1, \ldots, 1/\sigma_n). \tag{15}$$

- If none of the singular values are 0 and if $m \geq n$, then the left inverse of $\mathbf{A}$ is

$$\mathbf{V}\mathbf{D}^{-1}\mathbf{U}^T. \tag{16}$$

  We will extend this to the notion of the pseudo-inverse when we discuss least-squares methods.

- We can use the SVD to prove that any square matrix **A** may be decomposed

$$\mathbf{A} = \mathbf{R}\mathbf{S}. \tag{17}$$

  where **R** is orthogonal and **S** is symmetric, positive definite.

- The computation required to compute the SVD is $O(m^2 n)$, but less is required when **U** is not needed.

Practice Problems / Potential Test Questions

1. Prove that if the eigenvalues of a symmetric matrix $\mathbf{A}$ are not unique then the spectral decomposition is not unique. Explain the cause of the non-uniqueness.

2. Suppose that $\mathbf{A}$ is a symmetric matrix such that $\mathbf{AA} = \mathbf{A}$. Describe $\mathbf{A}$ as precisely as you can.

3. A skew-symmetric matrix is one where $\mathbf{A} = -\mathbf{A}^\top$. Find the diagonal terms and the eigenvalues of $\mathbf{A}$.

4. Find the determinant and the eigenvalues of a symmetric, positive definite matrix $\mathbf{A}$ from its Cholesky factorization.

5. Consider $m \times n$ matrix $\mathbf{A}$, with $m \geq n$, and vector $\mathbf{y} \in \mathbb{R}^m$. Use the SVD of $\mathbf{A}$ to show that the smallest magnitude $\mathbf{x}$ such that $\mathbf{Ax} = \mathbf{y}$ is entirely in the row-space of $\mathbf{A}$.

6. Show that the two forms of the SVD (10) and (11) are equivalent.

Problems For Grading

Submit solutions to the following problems on Monday, February 13th, as part of HW 4.

1. (**20 points**) Consider the singular value decomposition $\mathbf{A} = \mathbf{UDV}^T$, with

$$\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n),$$

$$\mathbf{D} = \text{diag}(\sigma_1, \ldots, \sigma_n),$$

and

$$\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n).$$

Suppose also that the last $n-k$ singular values are 0. Show that $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ form a basis for the nullspace of $\mathbf{A}$. Hint: think through carefully what you need to prove here since there are several parts to the proof.

2. (**15 points**) Show that if all eigenvalues of a symmetric $n \times n$ matrix $\mathbf{A}$ are non-negative then for every vector $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x}^\top \mathbf{Ax} \geq 0.$$