GAUSS Seminar: SVD, PCA, …

3/24/2012

Xin Li

Recall: Represent Matrix Using Its Column Basis

- Assume $D \in R^{m \times n}$ and m and n are a "large" number (like 10^6).

D is a "vector of vector" $= [d_1 \;\; d_2 \;\; \dots \;\; d_n]$, $d_j \in R^m$

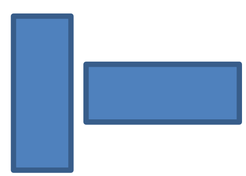
- How many entries in D?
- Answer: $m \times n$
- Takes a lot of memory to store D (in computer)
- What if we know that the columns of D come from a much smaller k dimensional subspace of R^m? Say $k = 6$.

$D = [d_1 \;\; d_2 \;\; \dots \;\; d_n]$ should have rank $\leq k$

- We can use this to "identify" each column of D.

=reconstruct each column from these k basis vectors by using the k coefficients in the linear combination: $d_j = v_{j1}\vec{\boldsymbol{u}}_1 + v_{j2}\vec{\boldsymbol{u}}_2 + \cdots + v_{j6}\vec{\boldsymbol{u}}_6$

Recall: Dimension Reduction

- Let $U = [\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2, \dots, \vec{\boldsymbol{u}}_6]$ be the matrix formed by a basis of the column vectors of D.
- So, there is a matrix V such that
$D = [d_1, d_2, \dots, d_n] = [\vec{\boldsymbol{u}}_1, \vec{\boldsymbol{u}}_2, \dots, \vec{\boldsymbol{u}}_6] \begin{bmatrix} v_{11} & v_{12} & v_{1n} \\ v_{21} & v_{22} & \dots v_{2n} \\ \vdots & \vdots & \vdots \\ v_{61} & v_{62} & v_{6n} \end{bmatrix}$
- What is the dimension of V?
- Answer: $6 \times n$
- How to find U and V?
- Note that $D = UV$= 

How to find (tall matrix) U and (wide matrix) V ?

- This can be done in many ways
- We will use the one based on SVD
- SVD=Singular Value Decomposition
- SVD is a universal factorization method of matrices
- This means that every matrix has its own SVD
- **SVD Theorem**: Let $D \in R^{m \times n}$. Then there exist orthogonal matrices $P \in R^{m \times m}$, $Q \in R^{n \times n}$, and diagonal matrix $\Sigma \in R^{m \times n}$ with nonnegative entries such that
$$D = P\Sigma Q^t.$$

How to find (tall matrix) U and (wide matrix) V ? (II)

- If D has rank k, then only k entries on the diagonal of Σ are positive.
- Write $\Sigma = \Sigma^{1/2}\,\Sigma^{1/2}$ ($m = n$, taking the square roots)
- Then $U = P\Sigma^{1/2}$ and $V = \Sigma^{1/2}Q^t$
- Check they have the desired properties
- Not unique: a different choice is

$$U = [\vec{\boldsymbol{p}}_1, \vec{\boldsymbol{p}}_2, \dots, \vec{\boldsymbol{p}}_k] \text{ and } V = \Sigma Q^t$$

- Indeed, $\vec{\boldsymbol{p}}_1, \dots, \vec{\boldsymbol{p}}_k$ are called principle vectors

Deriving SVD

- How to find the desired P, Q, Σ in $D = P\Sigma Q^t$?
- “reverse engineering”
- Assume that we have them and $D = P\Sigma Q^t$
- Then note that $D^t = (P\Sigma Q^t)^t =$
- $= Q\Sigma^t P^t$
- So, $D^t D = Q\Sigma^t P^t P\Sigma Q^t =$
- $= Q(\Sigma^t\Sigma)Q^t$
- Note that $\Sigma^t\Sigma = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots \end{bmatrix} \in R^{n\times n}$

We found Σ and Q!

- The diagonal entries σ_i of Σ have the property that $\sigma_i^2, i = 1,2, \dots, n$, are the eigenvalues of the (semi-positive) symmetric matrix $D^t D$
- Q has columns that are eigenvectors of $D^t D$
- What about P? (HW: Verify the following)
- The diagonal entries σ_i of Σ have the property that $\sigma_i^2, i = 1,2, \dots, n$, are the eigenvalues of the (semi-positive) symmetric matrix DD^t
- P has columns that are eigenvectors of DD^t

Quiz

- Recall that $D \in R^{m \times n}$.
- How many eigenvalues (counting the multiplicity) does $D^t D$ have?
- Answer: n
- How many eigenvalues (counting the multiplicity) does DD^t have?
- Answer: m
- True or false: Since σ_i^2 are eigenvalues of both $D^t D$ and DD^t, they must have the same eigenvalues.
- If D has rank k,
  - (i) what is the rank of DD^t? $D^t D$?
  - (ii) how many σ_i are nonzero?

Proof of SVD

- Consider matrix $A = D^t D \in R^{n \times n}$
- Then A has only nonnegative eigenvalues (semi-positivity)
- Denote the eigenvalues of A by $\lambda_i(A)$, $i = 1,2, \ldots, n$.
- There are n orthogonal eigenvacotrs $\vec{\boldsymbol{q}}_i \in R^n$such that

$$A\vec{\boldsymbol{q}}_i = \lambda_i(A)\vec{\boldsymbol{q}}_i, \qquad i = 1,2, \ldots, n$$

- Let $Q = [\vec{\boldsymbol{q}}_1, \vec{\boldsymbol{q}}_2, \ldots, \vec{\boldsymbol{q}}_n]$ and

$$\sigma_i = \sqrt{\lambda_i(A)} \geq 0, \qquad i = 1,2, \ldots, n$$

- Let $\Sigma = diag(\sigma_1, \ldots, \sigma_n, \text{pad with 0 if needed})$

Will These Choices for Σ and Q work?

- Let

$$\vec{p}_i = \frac{D\vec{q}_i}{\sqrt{\lambda_i(A)}}, i = 1,2, \dots, k$$

- Magical moment:

$$DD^t\vec{p}_i = \lambda_i(A)\vec{p}_i \text{ (Verify this!)}$$

- Let

$$P = [\vec{p}_1, \vec{p}_2, \dots, \vec{p}_k] = \left[\frac{D\vec{q}_1}{\sqrt{\lambda_1(A)}}, \frac{D\vec{q}_2}{\sqrt{\lambda_2(A)}}, \dots, \frac{D\vec{q}_n}{\sqrt{\lambda_n(A)}}\right]$$

- $= DQ\Sigma^{-1}$.
- So, $P\Sigma = DQ$. Thus $D = P\Sigma Q^t$

Almost Done

- Are we done?
- Not yet.
- P is not a square (only k columns are defined, still need $m-k$ columns).
- How?
- Easy: just add $m-k$ orthogonal vectors from the subspace orthogonal to the first n columns.
- DONE

Use of SVD for PCA Algorithm

- PCA=principle component analysis
- Short answer: sort σ_i from largest to smallest and use $\vec{p}_1, \dots, \vec{p}_{k_0}$ ($k_0 \leq k$) as the "principle components" in representing all other columns of D.
- σ_i kind of measures the "importance" of $\vec{p}_i$
- $D = P\Sigma Q^t \approx \left[\vec{p}_1, \vec{p}_2, \dots, \vec{p}_{k_0}\right] V$
- Each column of D is coded by the k_0 dimensional columns of V.

Assignments

- Do all the "Verify" in the slides and write up your solutions. I will collect them next Friday.
- Keep using Matlab and read the documentation about eig, svd, …