More on SVD

Monday, October 10, 2005 11:04 AM

The basic idea of SVD is that, given any rectangular matrix A with dimensions $m \times n$,

 $\begin{array}{l} A = U \ \Sigma \ V^{T}; \\ A: \ m \times n, \ m \ge n \\ \Sigma: \ m \times n \\ V: \ n \times n \\ U^{T}U = I_{m}, \ V^{T}V = I_{n} \end{array}$

$$\Xi = \begin{pmatrix} \overline{\delta} & 0 \\ 0 & \overline{\delta} \\ 0 & \overline{\delta} \end{pmatrix}_{MY} M$$

$$Rank(A) = r$$

Last time we looked at a matrix called \widetilde{A}

$$\widehat{A} = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \text{ where } \widehat{A}(A) = \pm \sigma(A)$$

Observe that \tilde{A} is symmetric so its eigenvectors \boldsymbol{z} can be normalized

 $\langle \cdot, \cdot \rangle$

 $||_{z_i}||_{2^{-1}}$

10.10 D

10.10 Page 1

$$\frac{1}{2} = \left(\begin{array}{c} \vec{x} \\ \vec{y} \end{array} \right)_{1} = \left(\begin{array}{c} \vec{x} \\ -\vec{y} \end{array} \right)_{1} = \left(\begin{array}{c} \vec{y} \end{array} \right)_{1} = \left(\begin{array}{c} \vec{x} \\ -\vec{y} \end{array} \right)_{1} = \left(\begin{array}{c} \vec{y} \end{array} \right)_{1} = \left(\begin{array}(c) \vec{y} \end{array} \right)_{1} = \left(\begin{array}(c) \vec{y} \end{array} \right)_{1} = \left($$

is also an eigenvector

Thus the mass is distributed over both parts of the eigenvector, i.e. $x^T x = 1/2$, $y^T y = 1/2$

Define a new matrix,

$$\widetilde{Z} = \begin{pmatrix} \chi & \chi \\ \gamma & -\gamma \end{pmatrix} \overline{f_2} \quad for \quad \delta \neq 0$$

Then we claim that \tilde{A} can be written as

$$\begin{split} \tilde{A} &= \tilde{z} \wedge \tilde{z}^{\mathsf{T}} = \begin{pmatrix} \mathbf{x} & \mathbf{x} \\ \mathbf{y} & \mathbf{y} \end{pmatrix} \begin{pmatrix} \mathbf{z} & \mathbf{r} \\ \mathbf{\theta} & \mathbf{z} \end{pmatrix} \tilde{z}^{\mathsf{T}} \\ \tilde{A} &= \begin{pmatrix} \mathbf{\theta} & A \\ A^{\mathsf{T}} & \mathbf{\theta} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{\theta} & \lambda \mathbf{z} & \mathbf{y}^{\mathsf{T}} \\ \mathbf{y} \mathbf{z} & \mathbf{x}^{\mathsf{T}} & \mathbf{\theta} \end{pmatrix} \\ &= \tilde{A} \colon \chi \mathbf{z}, \, \tilde{y}^{\mathsf{T}} \end{split}$$

Jordan Canonical Form

Monday, October 10, 2005 11:17 AM

JCF is more useful as an analytical tool than numerical

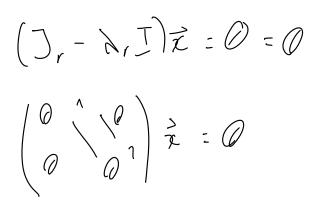
Every matrix A can be written as $A = QJQ^{-1}$ -where

$$\exists = \left(\begin{array}{c} J_1 & 0 \\ 0 & J_k \end{array} \right)$$

Where

$$\sum_{r}^{=} \left(\begin{array}{c} \partial r & & 0 \\ & & & 1 \\ 0 & & \lambda r \end{array} \right)$$

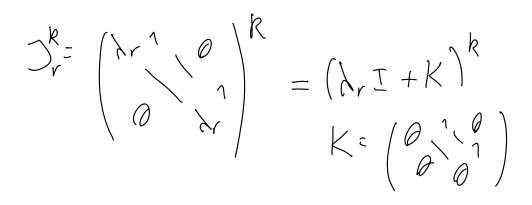
The rank of $J_r - \lambda_r I$?



$$(\mathbf{D}_r - \mathbf{\lambda}_r \mathbf{t})_{e_1} = 0$$

e1 is an eigenvector of this matrix

JCF is a useful for computing powers of matrices: $A^2 = QJQ^{-1}QJQ^{-1} = QJ^2Q^{-1}$ and $A^n = QJ^nQ^{-1}$

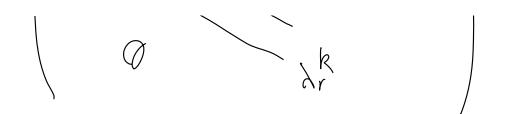


So now, we want to look at

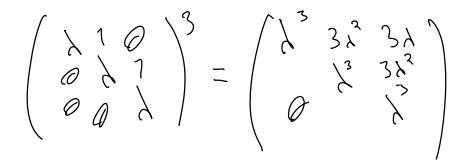
$$(\lambda_{r} I + K)^{k} = \sum_{j=0}^{k} \binom{k}{j} \lambda_{r}^{k-j} K^{j}$$

$$= \left(\lambda_{r}^{k} \binom{k}{1} \lambda_{r}^{k-1} - \binom{k}{n_{r}-1} \lambda_{r}^{k-(n_{r}-1)} \right)$$

$$= \left(\lambda_{r}^{k} \binom{k}{1} \lambda_{r}^{k-1} - \binom{k}{n_{r}-1} \lambda_{r}^{k-(n_{r}-1)} \right)$$

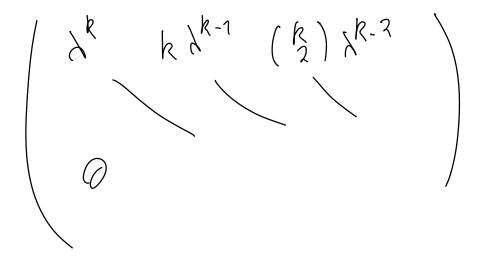


Jordon Canonical form



This is important in doing powers of matrices; if we have $|\lambda| < 1$

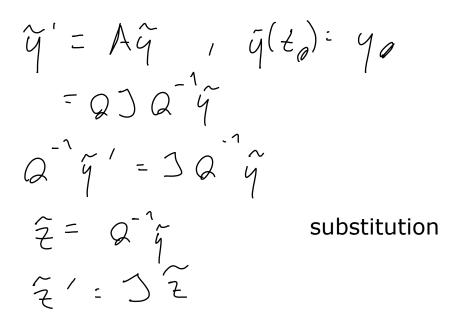
And raise that matrix to the *k*th power, we'd get



Which approaches 0, but not in a regular way -- for example the terms $k\lambda^{-1}$ go up, then down

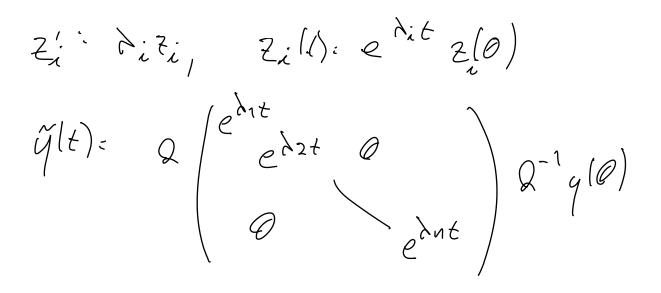
Monday, October 10, 2005 11:28 AM

Functions of matrices:



Say J is a diagonal matrix





 $\equiv e^{At} \gamma (Q)$ $F(A): QF(J)Q^{-1}$

Important Theorem

Monday, October 10, 2005 11:33 AM

is called the characteristic equation; the Cayley-Hamilton theorem says that

$$\Psi(A): \ \alpha_{0}I + - + \alpha_{m}A^{m} = 0$$

Every matrix satisfies this!

Alternatively,

$$\alpha_{0}A^{-1} + \alpha_{1}I + \alpha_{2}A + - + \alpha_{m}A^{n-1} = 0$$

$$A^{-1} = p_{1} \pm p_{2}(A) + - + \beta_{n-1}A^{n-1}$$

$$A^{-1} = Polynomial of degree n-1 in A$$

The amazing thing is that this says that any analytic function F(A) can be expressed by polynomials (require "analytic" so that a power series construction exists)! That is, if

$$\alpha_{1}I + \alpha_{n}A + - + \alpha_{m}A^{2} = 0$$

Then $A^n = poly(A)$ of degree n-1

$$A^{n_{12}}: \gamma_{0} I + \gamma_{1} A + - + \gamma_{n_{12}} A^{n_{12}}$$

$$= A \cdot A^{n_{12}}$$

$$= A \cdot (degree n \cdot 1)$$

For example if we have a 2×2 matrix, any power of it can be written down as a polynomial of degree 2

More on SVD

The SVD is called "robust" by engineers, but JCF is called "dynamic"

1) Given A, find \tilde{A} so that \tilde{A} is rank r and $||A-\tilde{A}||_2$ is minimal; in other words, what is the best rank r approximation?

Answer: if
$$A: \mathcal{U} \in \mathcal{V}^T$$
, then
 $\widetilde{A}: \mathcal{U} \cap \mathcal{V}^T$, $\Omega: \begin{pmatrix} \sigma_1 & \sigma_2 \\ \sigma_1 & \sigma_2 \end{pmatrix}$

i.e. compute SVD, take the first r values, and throw the rest away

from last time

$$= \| (\xi - \Omega) \|_{F}$$

$$= (\sigma_{1}, \sigma_{r})^{1/2}$$

$$\Omega = (\sigma_{1}, \sigma_{r})^{1/2}$$

We talked before about

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} , \text{ the best rank 1 approx is } \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Aside: Say we have $A = \tilde{A} + E$ with $||E||_F = \varepsilon$

Find \widetilde{A} so that $\mathbb{A} \xrightarrow{\sim} \widetilde{A} \stackrel{\|}{\prec} \stackrel{\mathcal{L}}{\leftarrow}$ and rank \widetilde{A} is minimum

Say *A* has singular values $\sigma_1, \dots, \sigma_n$ and find lowest *r* so that $\sigma_1 \neq \cdots = t \sigma_r \leq \varepsilon$ Say we have a matrix A, and we want it to be orthogonal, but it isn't. The singular values of an orthogonal matrix are $\frac{1}{2}$ How to get the best orthogonal matrix approximation to A?

Say we have two matrices containing data

They contain the same data, maybe perturbed by a rotation

$$\frac{\min|||A - ||Q||}{2^{\tau_0:t}} = \min||A - ||Q||$$

What's the orthogonal Q best for this?

This is only known for the Frobenius norm; for other norms nobody knows what Q is

Monday, October 10, 2005 11:56 AM

Polar decomposition

$$\mathcal{R} = \mathbf{a} + i\mathbf{b}$$
$$= \mathbf{p} e^{i\mathbf{0}} \qquad \mathbf{p} = \mathbf{p} e^{i\mathbf{0}}$$

A is an $n \times n$ matrix with an SVD $U \Sigma V^T$

rewrite this as

UVI VZVI

= QH, H is hermitian, Q orthogonal

Claim that this is the polar decomposition, an orthogonal matrix times a hermitian

we already saw that $\|b A_{\mathcal{L}}\|_{\mathcal{I}_{\mathcal{I}}}$ is solved by $\neq A^{\dagger}b$

But maybe we "don't know" A

A=A+E, ILElla SE

Say $\widehat{\mathbb{A}}$ is our approximation, then

Perturbation Theory

If for some natural norm, ||A|| < 1, we have

1)
$$\underline{T} - A$$
 is non-singular
2) $\| (\underline{T} - A)^{-1} \| \leq \frac{1}{1 - \| A \|}$
 $\leq \frac{1}{1 + \| A \|}$

() Suppose *A* is Singular, then there is a vector \mathbf{z} so that $A\mathbf{z} = 0$

If (I-A) is singular, there exists nonzero z with (I-A)z = 0

so $z: A^{2} \implies ||A|| > 1$, a contradiction

Proof of (2)

Know inverse exists, so write
$$J = (J - A)(J - A)^{2}$$

$$\|J = \|(J - A)(J - A)^{2}\| \leq \|(J - A)\| \cdot \|(J - A)^{2}\|$$

$$\leq (I + \|A\|)(\leq \mathbb{O})$$
So
$$\frac{1}{1 + |A|} \leq \|(J - A)^{2}\|$$

we had $(\underline{\tau} - A)^{-1} (\underline{\Gamma} - A) = \underline{T}$ $(\underline{\tau} - A)^{-1} - (\underline{\tau} - A)^{-1} A = \underline{T}$ $(\underline{\tau} - A)^{-1} = \underline{\tau} + (\underline{\tau} - A)^{-1} A$