MATH 612
Computational methods for equation solving and function minimization – Week # 2

Instructor: Francisco-Javier ‘Pancho’ Sayas

Spring 2014 – University of Delaware

Plan for this week

- Discuss any problems you couldn't solve of Lectures 1 and 2 (Lectures are the chapters of the book)
- Read Lectures 3, 4, and 5
- I will talk about Lectures 3, 4, and 5, but will not cover everything in the book, and especially, not in the same order. I will respect notation though.
- The first HW assignment will be given. It includes problems from Lectures 1–4.

Warning

Typically, I'll keep on updating, correcting, and modifying the slides until the end of each week.

NORMS

Norms for vectors

A norm in $\mathbb{C}^n$ is a map $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ satisfying

- Positivity: $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$
- The triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$
- Homogeneity: $\|\alpha x\| = |\alpha| \, \|x\|$

These properties hold for arbitrary $x, y \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$.

Important examples

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} \qquad 1 \leq p < \infty, \qquad \|x\|_\infty = \max_{j=1,\ldots,n} |x_j|.$$

The values $p = 1$ and $p = 2$ are the most commonly used.

Other norms

If W is an invertible matrix, then

$$\|Wx\|_p \qquad 1 \le p \le \infty$$

are also norms (W as in weight). Think for instance of a diagonal matrix $W = \text{diag}(w_1, \dots, w_n)$ and $p = 2$

$$\left(\sum_{j=1}^{n} |w_j|^2 |x_j|^2\right)^{1/2},$$

and with $p = 1$

$$\sum_{j=1}^{n} |w_j||x_j|.$$

Several MATLAB formulas

To compute the 2-norm you can use all of these strategies:

```
>> x=[3 1 2 5 6]';     % column vector
>> sqrt(sum(abs(x).^2))
ans =
   8.660254037844387
>> sqrt(dot(x,x));     % results ommited
>> sqrt(x'*x);     % only works if x is a column vector
>> norm(x,2);
>> norm(x);       % the 2-norm is the default norm
```

P.S.

If there's a Matlab function for something, use it (unless instructed to DIY)

Continues...

```
>> x=[3 1 2 5 6]';     % column vector
>> sum(abs(x))
ans =
    17
>> norm(x,1)
ans =
    17
>> max(abs(x))
ans =
     6
>> norm(x,Inf)
ans =
     6
```

A three line experiment

Find numerical evidence (using a single vector) that the $p-$norm of a vector is a continuous function of the parameter p with correct limit as $p \to \infty$.

The shape of the p-balls

The sets

$$\{x \in \mathbb{R}^2 \ : \ \|x\|_p \leq 1\}$$

have very different shapes for $p = 2$ (circle), $p = \infty$ (square parallel to the axes) $p = 1$ (square parallel to the main diagonals).

Note that

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1.$$

(Can you prove this? Can you see why this proves that for the same radius, the ∞ balls are larger than the 2 balls, and both are larger than the 1 balls?)

Operator (or induced) norms for matrices

For any $A \in \mathbb{C}^{m,n}$ and arbitrary p we can define the norm

$$\|A\|_{(p)} = \sup_{0 \neq x \in \mathbb{C}^n} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{x \in \mathbb{C}^n, \|x\|_p = 1} \|Ax\|_p.$$

(The supremum is actually a maximum.) These are actually norms valid in the spaces of matrices for all sizes. They additionally satisfy:

1. By definition... $\|Ax\|_p \leq \|A\|_{(p)} \|x\|_p$
2. By a simple argument... $\|AB\|_{(p)} \leq \|A\|_{(p)} \|B\|_{(p)}$
3. By a simpler argument... $\|I\|_{(p)} = 1$.

The Frobenius norm

$$\|A\|_F = \Big(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \Big)^{1/2} = \sqrt{\text{trace}(A^* A)}$$

satisfies the second property but not the third.

Computing the norms

The Frobenius norm is easy to compute (it's just a formula). There are formulas for $\|A\|_{(1)}$ and $\|A\|_{(\infty)}$. The computation of $\|A\|_{(2)}$ will be shown later this week. (It requires computing one eigenvalue.)

Formulas (the proofs are in the book)

$$\|A\|_{(1)} = \max_j \|a_j\|_1 \qquad \text{where } a_j \text{ are the columns of } A$$

$$\|A\|_{(\infty)} = \max_i \|a_i^*\|_1 \qquad \text{where } a_i^* \text{ are the rows of } A$$

Note that

$$\|A\|_{(\infty)} = \|A^*\|_{(1)}$$

More Matlab...

```
>> A=[-1 -2 3;4 -5 6]
A =
    -1    -2     3
     4    -5     6
>> norm(A,1)
ans =
     9
>> max(sum(abs(A)))     % sum adds by columns (default)
ans =
     9
>> norm(A,Inf)
ans =
    15
>> max(sum(abs(A')))
ans =
    15
>> max(sum(abs(A),2))
ans =
    15
```

Now you think

- Prove that if Q is a unitary, then

$$\|QA\|_F = \|A\|_F \qquad \text{and} \qquad \|QA\|_{(2)} = \|A\|_{(2)}$$

- What is $\|I\|_F$? (Hint. It depends on the dimension.)
- What is computed with the following Matlab line?

```
sum(abs(A(:).^2))
```

  More importantly, how?
- Figure out what is the Matlab command for the Frobenius norm of a matrix.

The SVD

Important announcement

This is one of the most important concepts in mathematics you will ever learn. It is used with different names in all kinds of areas of mathematics, signal processing, statistics,...

SVD stands for singular value decomposition

This lecture is quite theoretical. You'll need to be patient and alert!

The matrix A^*A

Let A be a complex $m \times n$ matrix. Consider the $n \times n$ matrix A^*A.

- It is Hermitian ($(A^*A)^* = A^*A$). Therefore its eigenvalues are real, it is diagonalizable, and its eigenvectors can be made build an orthonormal basis of $\mathbb{C}^n$ ($\mathbb{R}^n$ if A is real).
- It is positive semidefinite:

  $$x^*(A^*Ax) = (x^*A^*)(Ax) = (Ax)^*(Ax) = \|Ax\|^2 \geq 0 \quad \forall x \in \mathbb{C}^n.$$

  Therefore its eigenvalues are non-negative (real) numbers.
- Its nullspace is the same as the one for A:

  $$Ax = 0 \qquad \Longleftrightarrow \qquad A^*Ax = 0.$$

  (See the previous item for the proof.) Therefore, by the rank-nullity theorem A^*A has the same rank as A.

Applying the spectral theorem

We organize the eigenvalues of A^*A in decreasing order:

$$\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_r^2 > \sigma_{r+1}^2 = \ldots = \sigma_n^2 = 0.$$

We choose an orthonormal basis of eigenvectors

$$A^*Av_j = \sigma_j^2 v_j \qquad \text{null}(A) = \text{nul}(A^*A) = \langle v_{r+1}, \ldots, v_n \rangle.$$

We then consider the following vectors

$$u_j = \frac{1}{\sigma_j} A v_j, \qquad j = 1, \ldots, r.$$

Elementary, my dear Watson

The vectors u_j are orthonormal

$$\begin{aligned} u_i^* u_j &= \left(\frac{1}{\sigma_i} A v_i\right)^* \left(\frac{1}{\sigma_j} A v_j\right) = \frac{1}{\sigma_i \sigma_j} v_i^* A^* A v_j \\ &= \frac{\sigma_j^2}{\sigma_i \sigma_j} v_i^* v_j = \delta_{ij} \end{aligned}$$

and they span the range of A. Moreover, for all x

$$\begin{aligned} Ax &= A \underbrace{\left(\sum_{j=1}^{n} (v_j^* x) v_j \right)}_{=x} = \sum_{j=1}^{n} (v_j^* x) A v_j \\ &= \sum_{j=1}^{\color{red}{r}} (v_j^* x) \sigma_j u_j = \sum_{j=1}^{r} \sigma_j (v_j^* x) u_j. \end{aligned}$$

The reduced SVD

The reduced SVD is the matrix form of the equality

$$Ax = \sum_{j=1}^{r} \sigma_j (v_j^* x) u_j.$$

$$A = \widehat{U}\widehat{\Sigma}\widehat{V}^*$$

- $\widehat{V}$ is the $n \times r$ matrix whose columns are the orthonormal vectors $\{v_1, \ldots, v_r\}$.
- $\widehat{\Sigma}$ is the $r \times r$ diagonal matrix whose diagonal entries are the positive numbers

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0.$$

- $\widehat{U}$ is the $m \times r$ matrix whose columns are the orthonormal vectors $\{u_1, \ldots, u_r\}$.

Terminology

$$Ax = \sum_{j=1}^{r} \sigma_j (v_j^* x) u_j \qquad A = \widehat{U}\widehat{\Sigma}\widehat{V}^*$$

- The numbers

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0 \qquad r \text{ is the rank of } A$$

  are the singular values of A.
- The orthonormal vectors $\{v_1, \ldots, v_r\}$ are the right singular vectors. (They are the eigenvectors of A^*A except those corresponding to the zero eigenvalue.)
- The orthonormal vectors $\{u_1, \ldots, u_r\}$ are the left singular vectors.

Bear with me

There's a lot of information encoded in this expressions:

$$Ax = \sum_{j=1}^{r} \sigma_j (v_j^* x) u_j \qquad A = \widehat{U}\widehat{\Sigma}\widehat{V}^*.$$

Forget how we got it. Assume that we have just been told that the SVD exists. (Remember the rules on the three matrices.) Then:

$$A^* y = \sum_{j=1}^{r} \sigma_j (u_j^* y) v_j \qquad A^* = \widehat{V}\widehat{\Sigma}\widehat{U}^*.$$

$$A^* A x = \sum_{j=1}^{r} \sigma_j^2 (v_j^* x) v_j \qquad A^* A = \widehat{V}\widehat{\Sigma}^2 \widehat{V}^*$$

$$A A^* y = \sum_{j=1}^{r} \sigma_j^2 (u_j^* y) u_j \qquad A A^* = \widehat{U}\widehat{\Sigma}^2 \widehat{U}^*$$

All of this is true: prove it

Let

$$A = \widehat{U}\widehat{\Sigma}\widehat{V}^*$$

where the column vectors of $\widehat{U}$ and $\widehat{V}$ are orthonormal, and where $\widehat{\Sigma}$ is an $r \times r$ positive diagonal matrix with entires of non-increasing order. Then:

$$A^*Av_j = \sigma_j^2 v_j, \qquad AA^*u_j = \sigma_j^2 u_j,$$

$$Av_j = \sigma_j u_j, \qquad A^*u_j = \sigma_j v_j,$$

the rank of A is r and

$$\langle v_1, \ldots, v_r \rangle = \text{null}(A)^\perp = \text{range}(A^*)$$

$$\langle u_1, \ldots, u_r \rangle = \text{range}(A) = \text{null}(A^*)^\perp.$$

A geometric/information approach

Let A be an $n \times n$ real invertible matrix, with SVD[1]

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

By construction, the matrices U and V are unitary and can be taken to be real.
Let $x \in \mathbb{R}^n$ be such that

$$\|x\|_2^2 = 1 = x_1^2 + \ldots + x_n^2.$$

We then decompose ($c = V^* x$)

$$x = c_1 v_1 + \ldots + c_n v_n = (v_1^* x) v_1 + \ldots + (v_n^* x) v_n,$$

where

$$\|c\|_2^2 = c_1^2 + \ldots + c_n^2 = 1.$$

[1] we eliminate the hats, because in this case the reduced SVD is equal to the full SVD

A geometric/information approach (2)

Then

$$Ax = \sigma_1 c_1 u_1 + \ldots + \sigma_n c_n u_n = y_1 u_1 + \ldots + y_n u_n,$$

where

$$\left(\frac{y_1}{\sigma_1}\right)^2 + \ldots + \left(\frac{y_n}{\sigma_n}\right)^2 = 1.$$

In other words...

Let A be an invertible matrix with singular values

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0.$$

Then there exists an orthogonal reference frame in $\mathbb{R}^n$ such that the image of the unit ball by A is the hyperellipsoid

$$\left(\frac{y_1}{\sigma_1}\right)^2 + \ldots + \left(\frac{y_n}{\sigma_n}\right)^2 = 1.$$

The 2-norm of a matrix

We go back to the genera case

$$Ax = \sum_{j=1}^{r} \sigma_j (v_j^* x) u_j \qquad A = \widehat{U}\widehat{\Sigma}\widehat{V}^*.$$

Then

$$\|Ax\|_2^2 = \sum_{j=1}^{r} \sigma_j^2 |(v_j^* x)|^2 \le \sigma_1^2 \sum_{j=1}^{r} |(v_j^* x)|^2 \qquad (\sigma_1 \ge \sigma_j \quad \forall j)$$

$$\le \sigma_1^2 \sum_{j=1}^{n} |(v_j^* x)|^2 = \sigma_1^2 \|x\|_2^2$$

This proves (how?) that

$$\|A\|_{(2)} = \sup_{0 \ne x \in \mathbb{C}^n} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1 = \text{max singular value}$$

The full SVD

Again, for a general $m \times n$ matrix A,...

$$Ax = \sum_{j=1}^{r} \sigma_j (v_j^* x) u_j \qquad A = \widehat{U}\widehat{\Sigma}\widehat{V}^*.$$

Now:

- Build an orthonormal basis of $\mathbb{C}^n$, $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ by completing the right singular vectors with orthonormal vectors (from null(A)).
- Build an orthonormal basis of $\mathbb{C}^m$, $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$ by completing the left singular vectors with orthonormal vectors (from null(A^*)).

This is the same as adding $n - r$ columns to the right of $\widehat{V}$ and $m - r$ columns to the right of $\widehat{U}$ making the resulting matrices (V and U) unitary.

The full SVD (2)

We then create an $m \times n$ matrix (same size as A)

$$\Sigma = \begin{bmatrix} \widehat{\Sigma} & 0 \\ 0 & 0 \end{bmatrix}$$

by adding $n - r$ columns to the right and $m - r$ rows to the bottom of $\widehat{\Sigma}$. It is simple to see that

$$A = U\Sigma V^* = \widehat{U}\widehat{\Sigma}\widehat{V}^*.$$

Note sizes in the reduced decomposition

$$(m \times r)(r \times r)(r \times n)$$

and in the full decomposition

$$(m \times m)(m \times n)(n \times n)$$

The pseudoinverse

Recall the reduced SVD

Every $m \times n$ matrix A can be decomposed in the form

$$A = \widehat{U}\widehat{\Sigma}\widehat{V}^*$$

- $\widehat{U}$ is $m \times r$ with orthonormal columns
- $\widehat{\Sigma}$ is $r \times r$ diagonal with positive entries

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$$

- $\widehat{V}$ is $n \times r$ with orthonormal columns

The MATLAB command `[U,S,V]=svd(A)` returns the full SVD.

The (Moore-Penrose) pseudoinverse

If

$$A = \widehat{U}\widehat{\Sigma}\widehat{V}^*$$

we define

$$A^+ = \widehat{V}\widehat{\Sigma}^{-1}\widehat{U}^*.$$

Note the following:

- If A is invertible, then $A^+ = A^{-1}$. Why?
- For a general matrix A

$$x = A^+ b \qquad \Longrightarrow \qquad A^*Ax = A^*b$$

  (we'll talk about least-squares in due time)
- If $Ax = b$ has a unique solution (A has full rank by columns and b is in the range of A), then $x = A^+b$.

Wasting our time?

Let

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Can we compute its SVD? Can we do it without going through the diagonalization of A^*A?

$$A^*A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

Note that A has rank one. This is good news because $\widehat{\Sigma}$ is a 1×1 matrix. Now we cheat!

Wasting our time? (2)

First...

$$\text{null}(A) = \left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\rangle$$

so we can choose

$$\widehat{V} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}.$$

Then

$$\sigma_1 u_1 = Av_1, \qquad \|u_1\|_2 = 1,$$

which gives

$$\widehat{U} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \qquad \widehat{\Sigma} = \begin{bmatrix} \sqrt{6} \end{bmatrix}$$

Wasting our time? (... and 3)

This is the reduced SVD

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{6} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

and this is the pseudoinverse

$$A^+ = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/6 & 1/6 \\ 1/6 & 1/6 \\ 1/6 & 1/6 \end{bmatrix}$$

Matlab computes the pseudoinverse with the command `pinv`

You have a couple ol little challenging SVD to compute by hand in the book.

Fast now! Give me another SVD for A

Low-rank approximation

Ways to look at the SVD

The SVD

$$A = \widehat{U}\widehat{\Sigma}\widehat{V}^* \qquad \widehat{\Sigma} = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

can be read as

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^*.$$

Follow me now... We have written A, which is a matrix of rank r, as the sum of r matrices of rank 1. (Careful with this: if you add matrices of rank one you might end up with a matrix of less rank. The orthogonality of the vectors plays a key role here.)

A compressed version of A

Let

$$A_k := \sum_{j=1}^{k} \sigma_j u_j v_j^* = \widehat{U}_k \widehat{\Sigma}_k \widehat{V}_k^*.$$

Then A_k is a matrix of rank k. Its first k singular values, left singular vectors and right singular vectors coincide with those of A.

Now this is quite surprising...

$$A - A_k = \sum_{j=k+1}^{r} \sigma_j u_j v_j^*,$$

so the singular values of $A - A_k$ are

$$\sigma_{k+1} \geq \ldots \geq \sigma_r$$

and $\|A - A_k\|_{(2)} = \sigma_{k+1}$.

The low-rank approximation

In the book you have a proof (not complicated) of the following fact:

If B has rank $k < r$, then

$$\|A - B\|_{(2)} \geq \sigma_{k+1},$$

so we have found the best approximation of A in spectral norm $\| \cdot \|_{(2)}$ by matrices of rank k or less.

Now you are going to code this.

What to do

Create a function

```
B=lowrank(A,R)
```

such that given any matrix A and a number $0 < R < 1$ (a rate),

- computes its SVD (let Matlab do this, but remember that you get the full SVD)
- finds the lowest k such that

$$\sigma_1^2 + \ldots + \sigma_k^2 \geq R\,(\sigma_1^2 + \ldots + \sigma_r^2)$$

- then compresses A to rank k

For the code, it might be convenient to create a vector with the squares of the diagonal elements of Σ.

Sneak peek

Compute the full SVD of A (U, Σ, V)
D = vector with the square of the diagonal of Σ
energy=sum of elements of D
c=0;
for $k = 1$: length of D
    $c = c + D(k)$
    if $c \geq R \times$ energy
        leave the loop
    end
end
keep the first k columns of U
keep the first columns of V
keep the upper $k \times k$ block of Σ
build the low rank approximation