신호 처리를 위한 행렬 계산 Programming Assignment 2

20180490 이재헌

A. Code Implementation

1. norm_vec.m

   a. 0-norm

   It is the number of non-zero elements.
   Therefore, I counted the number of non-zero elements using the for loop.

```
y = 0;
for i=1:length(x)
    if x(i) ~= 0
        y = y + 1;
    end
end
```

   b. 1-norm

   It is the sum of absolute values of each element.

```
y = sum(abs(x));
```

   c. 2-norm

   It is the Euclidean norm.

```
y = 0;
for i=1:length(x)
    y = y + (x(i))^2;
end
y = sqrt(y);
```

   d. Inf-norm

   It is the maximum value among the absolute values of elements.

```
y = max(abs(x));
```

2. norm_mtx.m

   a. Frobenius norm

   It is the Euclidean norm.

```
y = 0;
for i=1:row_dim
    for j=1:col_dim
        y = y + x(i,j)^2;
    end
end
y = sqrt(y);
```

   b. 1-norm

   $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|$, (column-wise maximum absolute sum),

   이를 코드로 구현하면

```
y = max(sum(abs(x)));
```

c. 2-norm

$$\|A\|_2 \equiv \max_{\|x\|_2=1} \|Ax\|_2$$

For calculating the 2 norm, we use the ‘power iteration method’. I will explain why and how this method works in part B. The algorithm for this method is as follows.

```
N = 100;
v = rand(col_dim, 1);
for i = 1:N
    t = x.' * x * v;
    v = t / norm_vec(t, 2);
end
y = norm_vec(x * v, 2);
```

d. Inf-norm

$\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|$ (row-wise maximum absolute sum),

이를 코드로 구현하면

```
y = max(sum(abs(x),2));
```

3. cond_num.m

$$\kappa(A) = \|A^{-1}\| \, \|A\| \ge \|A^{-1}A\| = 1.$$

Since we have implemented the norm of a matrix for each norm, we can simply use these functions.

```
y = norm_mtx(x, 1)*norm_mtx(inv(x),1);
```

For other norms, we can just change the second parameter of norm_mtx to what we want.

B. Why and how does the power iteration method work?

Let the SVD for an m x n matrix be $A = U\Sigma V^T$. Then, $A^T * A = V\Sigma U^T U\Sigma V^T = V\Sigma^2 V^T$ is the eigenvalue decomposition of $A^T A$. We want to find v_1.

We can express a random vector t with the right singular vectors which are the basis.

$$t = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

Then, if we multiply $B \coloneqq A^T A$ k times to t, we get

$$B^k t = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n$$

$$Since\ B^k v_i = \lambda_i^k v \quad for\ all\ i$$

Assume $|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_n|$, we can approximate $B^k t \cong c_1 \lambda_1^k v_1$ if k is large enough.

After normalizing $B^k t$, we get the final result v_1 as we desired.

C. Discussion about <test 1>

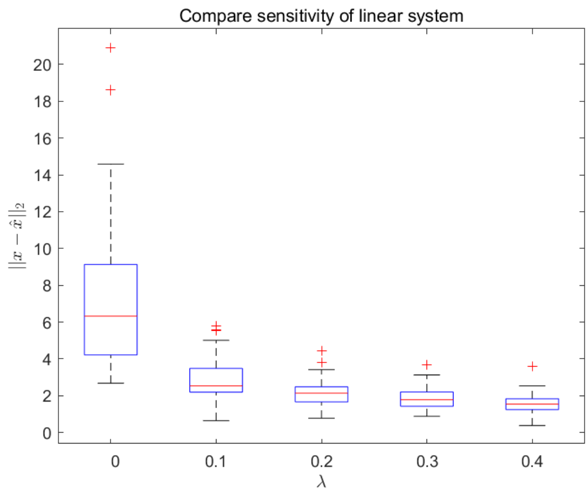
Here is the result of test 1.

Order	Custom_vector	Built_in_vector	Custom_matrix	Built_in_matrix	Custom_cond	Built_in_cond
"0"	6	6	NaN	NaN	NaN	NaN
"1"	20	20	23	23	29.144	29.144
"2"	9.5917	9.5917	17.416	17.416	18.796	18.796
"fro"	NaN	NaN	17.916	17.916	21.662	21.662
"inf"	7	7	19	19	27.873	27.873

As we can see from the table, the result of our customized functions coincides with the result of built-in functions. Therefore, our customized functions are verified.

D. Explain the result of <test 2> in terms of the condition number.

Here is the result of test 2.



Since ||x|| is fixed, we can see that the relative error decreases as λ increases.

I will discuss why this tendency happens below.

First, we should have some understanding on the condition number.

Q. Why is the condition number defined as $||A|| * ||A^{-1}||$?

Let e be the error in b. Assuming that A is a nonsingular matrix, the error in the solution $A^{-1}b$ is $A^{-1}e$. The ratio of the relative error in the solution to the relative error in b is

$$\frac{\|A^{-1}e\|}{\|A^{-1}b\|} \Big/ \frac{\|e\|}{\|b\|} = \frac{\|A^{-1}e\|}{\|e\|} \frac{\|b\|}{\|A^{-1}b\|}.$$

The maximum value (for nonzero b and e) is then seen to be the product of the two operator norms as follows:

$$\begin{aligned}\max_{e,b\neq 0}\left\{\frac{\|A^{-1}e\|}{\|e\|}\frac{\|b\|}{\|A^{-1}b\|}\right\} &= \max_{e\neq 0}\left\{\frac{\|A^{-1}e\|}{\|e\|}\right\}\max_{b\neq 0}\left\{\frac{\|b\|}{\|A^{-1}b\|}\right\} \\ &= \max_{e\neq 0}\left\{\frac{\|A^{-1}e\|}{\|e\|}\right\}\max_{x\neq 0}\left\{\frac{\|Ax\|}{\|x\|}\right\} \\ &= \|A^{-1}\|\ \|A\|.\end{aligned}$$

The 2–norm condition number represented with singular values is $\frac{\sigma_{max}(A)}{\sigma_{min}(A)}$.

Q. Represent $\kappa_2((A^H A + \lambda I)^{-1} A^H)$ in terms of $\sigma_{\max}(A), \sigma_{\min}(A)$ and λ, and then guess what happened to $\kappa_2((A^H A + \lambda I)^{-1} A^H)$ as λ increases. (hint: using SVD, $U\Sigma V^H = A$, $UU^H = U^H U = VV^H = V^H V = I$, and Σ is nonnegative real diagonal matrix)

$\kappa_2((A^H A + \lambda I)^{-1} A^H) = ?$ Let $A = U\Sigma V^H$

$A^H A + \lambda I = V\Sigma^H U^H U \Sigma V^H + \lambda I = V(\Sigma^2 + \lambda I)V^H$

$\therefore (A^H A + \lambda I)^{-1} = V(\Sigma^2 + \lambda I)^{-1} V^H$

$(A^H A + \lambda I)^{-1} A^H = V(\Sigma^2 + \lambda I)^{-1} V^H V \Sigma U^H = V(\Sigma^2 + \lambda I)^{-1} \Sigma U^H$

Since $\kappa_2(A) = \kappa_2(A^{-1})$,

$\kappa_2((A^H A + \lambda I)^{-1} A^H) = \kappa_2(U\Sigma^{-1}(\Sigma^2 + \lambda I)V^H)$

$$= \frac{\sigma_{max}(\Sigma^{-1}(\Sigma^2 + \lambda I))}{\sigma_{min}(\Sigma^{-1}(\Sigma^2 + \lambda I))} = k$$

$$\Sigma^{-1}(\Sigma^2 + \lambda I) = \begin{pmatrix} \frac{\sigma_1^2 + \lambda}{\sigma_1} & & & \\ & \frac{\sigma_2^2 + \lambda}{\sigma_2} & & \\ & & \ddots & \\ & & & \frac{\sigma_n^2 + \lambda}{\sigma_n} \end{pmatrix}$$

If $\sigma_{max} = \frac{\sigma_1^2 + \lambda}{\sigma_1}$ and $\sigma_{min} = \frac{\sigma_n^2 + \lambda}{\sigma_n}$,

$$k = \frac{\frac{\sigma_1^2 + \lambda}{\sigma_1}}{\frac{\sigma_n^2 + \lambda}{\sigma_n}} = \frac{\sigma_n}{\sigma_1} \cdot \left(\frac{\sigma_1^2 + \lambda}{\sigma_n^2 + \lambda}\right) = \frac{\sigma_n}{\sigma_1}\left(1 + \frac{\sigma_1^2 - \sigma_n^2}{\sigma_n^2 + \lambda}\right)$$

Therefore, if we assume σ_{max} and σ_{min} don't change as λ changes, k decreases as λ increases.

Although σ_{max} and σ_{min} depend on singular values and λ, if we assume σ_{max} and σ_{min} don't change as λ changes, the condition number decreases as λ increases.

For the case of our code, the singular values are as follows.

Singular values of A = 3.4166, 0.8617, 0.7890, 0.5745, 0.3422, 0.1543

Where A is a 6*6 random matrix.

- Discussion on the result of test 2

  As we can see the figure for the result, the relative error which is proportional to the condition number decreases as λ increases.

  We can check the condition numbers of A_pinv_reg for each λ.

λ	0	0.1	0.2	0.3	0.4
$\kappa(A_{pinv\ reg})$	22.1390	5.4314	3.7547	3.1954	2.7807
Maximum of the boxplot	14.5853	5.0131	3.4206	3.1318	2.5391
Median of the boxplot	6.3264	2.53	2.132	1.7891	1.5366

  We can see that the maximum of the boxplot follows the condition number. If we increase the number of repetition for our code, it will more closely follow the condition number.