Principles of Magnetic Resonance Imaging.
Exercise session 5

Exercise 1: Singular value decomposition and ill-conditioned matrix

A

Download the file `data.mat` and load it in matlab :

```
load 'data.mat'
```

Calculate the SVD of $\mathbf{A}$ and plot the singular values in logarithmic scale (use `semilogy`).
What do you notice? What is the rank of $\mathbf{A}$? Is $\mathbf{A}$ rank-deficient? What is the *numerical* rank of $\mathbf{A}$? What is the condition number of $\mathbf{A}$? Is $\mathbf{A}$ ill-conditioned?
For the solution of least squares problems, we should *somehow* compute $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. What happens if you try to compute $(\mathbf{A}^T\mathbf{A})^{-1}$ with matlab?
Plot the first 10 left singular vectors. Then the *last* 10. What do you notice?

B

The behavior of the singular vectors is typical in many ill-conditioned problems. Where are going to see what is the effect on the solution of the least squares problem

$$\arg\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \tag{1}$$

The solution to Eq. (1) can be explicitely given in terms of left/right singular vectors and singular values:

$$\mathbf{x}_{LS} = \sum_{i=1}^{n} \frac{\mathbf{u}_i^T\mathbf{b}}{\sigma_i}\mathbf{v}_i \tag{2}$$

Calculate $\mathbf{x}_{LS}$ and plot it together with the true solution which is part of the data, `x`.
As you can see, *there is no noise, but* $\mathbf{x}_{LS}$ *is completely wrong.*

C

To understand the reason of this discrepancy, let's see what happens to the coefficients in the sum. Plot on the same logarithmic graph, the values of σ_i and $|\mathbf{u}_i^T\mathbf{b}|$ for $i = 1, \ldots, 50$. Which terms decay faster?
Add also the plot of $\frac{|\mathbf{u}_i^T\mathbf{b}|}{\sigma_i}$. Comment what you see. It should be clear at this point that minimal rounding-off errors made by the computer (in the order of 10^{-15}) can have catastrophic effects. Also, note the type of perturbation in $\mathbf{x}_{LS}$ and how the singular vectors of $\mathbf{A}$ look like.

D

We will now investigate what happens when the right hand side (i.e. $\mathbf{b}$, the data) is perturbed by noise. Define the noisy signal, $\tilde{\mathbf{b}}$, by adding normally distributed noise to $\mathbf{b}$:

```
bt = b+10^(-6)*randn(size(b));
```

Plot $\mathbf{b}$ and $\tilde{\mathbf{b}}$, can you see the difference? Repeat the plots from part **C** for $\tilde{\mathbf{b}}$ instead of $\mathbf{b}$. What do you notice? Try also adding more noise: `bt = b+10^(-3)*randn(size(b));`

Exercise 2: Regularization

Truncated SVD

From the previous exercise, we learn that only the first component of the sum from Eq. (2) are useful. The perturbation in the data (noise) or the rounding-off errors are amplified by the last components, those corresponding to the smallest singular values.
Plot the *truncated SVD* solutions

$$\mathbf{x}_k = \sum_{i=1}^{k} \frac{\mathbf{u}_i^T \tilde{\mathbf{b}}}{\sigma_i} \mathbf{v}_i \tag{3}$$

for $k = 1, \dots, 20$ and `bt = b+10^(-3)*randn(size(b));`
Plot $\mathbf{x}_k$ together with the true solution, `x`. What do you notice? How many terms (k) do you need for a good solution? What is, approximately, the *cut-off value*, that is, α? What if k is too small? What if k is too large?

Tikhonov regularization

Plot the *Tikhonov* solutions

$$\mathbf{x}_\alpha = \sum_{i=1}^{n} \frac{\sigma_i \mathbf{u}_i^T \tilde{\mathbf{b}}}{\sigma_i^2 + \alpha} \mathbf{v}_i$$

for various values of α, together with the true solution. Try: $\alpha = 10^{-10}, 10^{-9}, \dots, 10^{-2}, 10^{-1}$.
The *Tikhonov* regularized solution can be written as

$$\mathbf{x}_\alpha = \arg\min_{\mathbf{x}} \{ \|\mathbf{A}\mathbf{x} - \tilde{\mathbf{b}}\|^2 + \alpha \|\mathbf{x}\|^2 \}$$

that is: a least-squares problem with a penalty on the (ℓ_2)-norm of the solution. The trade-off parameter α controls the balance between the solution norm $\|\mathbf{x}_\alpha\|$ and the residual norm $\|\mathbf{A}\mathbf{x}_\alpha - \tilde{\mathbf{b}}\|$
Plot in a logarithmic graph (use `loglog`) the points with coordinates $(\|\mathbf{A}\mathbf{x}_\alpha - \tilde{\mathbf{b}}\|, \|\mathbf{x}_\alpha\|)$ for several values of α and `bt = b+10^(-3)*randn(size(b));`
What do you notice? What is the best trade-off value of α?

This graph is called *L-curve* and it is used in practice to determine the optimal regularization parameter α.