Computer Science Department, Ben Gurion University of the Negev

Optimization methods with Applications - Spring 2020

Home assignment 1

Due on 5/4/2020.

March 22, 2020

1. **Vector norms, matrix norms, and inner products (Section 3 in NLA.pdf).**

   (a) We have learned that for any matrix $A \in \mathbb{R}^{m\times n}$, we have the induced norms

   $$\|A\|_1 = \max_{\mathbf{x}} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \max_j \left(\sum_{i=1}^{n} |a_{i,j}| \right),$$

   and similarly

   $$\|A\|_\infty = \max_{\mathbf{x}} \frac{\|A\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_i \left(\sum_{j=1}^{m} |a_{i,j}| \right).$$

   For the matrix

   $$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & -4 & 8 \\ -5 & 4 & 1 & 5 \\ 5 & 0 & -3 & -7 \end{bmatrix}$$

   Explicitly find a vector $\mathbf{x}$ that maximizes the terms above and fulfils the definitions. Find one for ℓ_1 and one for ℓ_∞. No need to prove your solution for $\mathbf{x}$ - just write an intuitive explanation.

   (b) Repeat the previos section for the ℓ_2 norm:

   $$\|A\|_2 = \max_{\mathbf{x}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sigma_{max}.$$

   where σ_{max} is the largest singular value of A (singular values are the square root of the eigenvalues of $A^T A$). Repeat the previous question for this norm as well (find $\mathbf{x}$ that maximizes the expression). It is recommended to use a computer program here, and submit the code.

2. **Least Squares (Section 5.1 in NLA.pdf)**

   (a) We are trying to approximate a vector $[x_1, x_2, x_3]$ by a constant c using ℓ_p norms. Assume $x_1 < x_2 < x_3$.

      i. Find the best approximation using c of the vector in ℓ_2 norm (Least squares):

      $$\min_c\{(c - x_1)^2 + (c - x_2)^2 + (c - x_3)^2\}.$$

      ii. Find the best approximation using c of the vector in ℓ_∞ norm:

      $$\min_c\{\max(|c - x_1|, |c - x_2|, |c - x_3|)\}.$$

      iii. Find the best approximation using c of the vector in ℓ_1 norm:

      $$\min_c\{|c - x_1| + |c - x_2| + |c - x_3|\}.$$

   Hint for (b) and (c): the solution is obtained by logic, not by calculations as in (a).

   (b) Find the best approximation in a least square sense for satisfying the linear system $A\mathbf{x} = \mathbf{b}$ where

   $$A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & -2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \quad , \mathbf{b} = \begin{bmatrix} 6 \\ 1 \\ 5 \\ 2 \end{bmatrix}.$$

   Write the normal equations and solve the problem using using the Cholesky factorization of $A^\top A$. You may use a computer and built-in functions for Cholesky, and provide the code.

   (c) Solve the least squares problem in subsection 2b using the (a) QR factorization of A, and (b) SVD factorization of A. Verify that you get the same answers. You may use a computer and built-in functions for QR, SVD.

   (d) Find the least squares solution of the system in (a), but now find a solution for which the first equation is almost exactly satisfied satisfied (let's say, such that $|r_1| < 10^{-3}$). Hint: use weighted least squares. You may solve the problem anyway you want.

3. QR **factorization**

   (a) Write two programs for getting the QR factorization using Gram Schmidt and Modified Gram Schmidt as presented in class.

   (b) Using the two codes you programmed, find the QR factorization of the matrix

   $$A = \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}$$

   for $\epsilon = 1, 1\text{e-}10$, for both algorithms (that is, perform four factorizations).

   (c) For all factorizations, calculate $\|Q^\top Q - I\|_F$. Which of the algorithms produces a better QR factorization? Explain.

4. **Regularized Least Squares and SVD** (Sections 5.4 - 5.5 in NLA.pdf.)

   For this question we assume that we have a full rank matrix $A \in \mathbb{R}^{m\times n}$ for which we have an SVD factorization $A = U\Sigma V^\top$, and denote the singular triplets as $(\mathbf{u}_i, \sigma_i, \mathbf{v}_i)$.

   We saw in class, that using the SVD, the solution of the LS problem $\arg\min_{\mathbf{x}}\{\|A\mathbf{x} - \mathbf{b}\|_2^2\}$, is given by

   $$\hat{\mathbf{x}} = (A^\top A)^{-1}A^\top\mathbf{b} = V\Sigma^{-1}U^\top\mathbf{b}.$$

   (a) (Not for submission) Show that $V\Sigma^{-1}U^\top = \sum_{i=1}^{\min(m,n)} \sigma_i^{-1}\mathbf{v}_i\mathbf{u}_i^\top$, where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the columns of U and V respectively.
   Guidance: this equality holds simply from the definition of matrix multiplication. Show that the (i, j) entry of both sides are identical.

   (b) Using the information above show:

   $$\hat{\mathbf{x}} = \sum_{i=1}^{\min(m,n)} \frac{1}{\sigma_i}(\mathbf{u}_i^\top\mathbf{b})\mathbf{v}_i.$$

   (c) Consider a full SVD, where $U \in \mathbb{R}^{m\times m}$ so we have m vectors $\mathbf{u}_i$ that span the whole R^m space. Assume that $\mathbf{b} = \sum_{i=1,..,m} \alpha_i\mathbf{u}_i$ for some α_i. Show that

   $$\hat{\mathbf{x}} = \sum_{i=1}^{\min(m,n)} \frac{1}{\sigma_i}\alpha_i\mathbf{v}_i.$$

(d) Now we consider the regularized LS problem $\arg\min\{\|A\mathbf{x}-\mathbf{b}\|_2^2+\lambda\|\mathbf{x}\|_2^2\}$. Show that the solution of this problem is given by the solution of the system (look 5.1.3 in NLA.pdf)

$$(A^\top A+\lambda I)\mathbf{x}=A^\top\mathbf{b}.$$

Also, show that the matrix $(A^\top A+\lambda I)$ is always positive definite for $\lambda>0$, even if A is not full rank (use the definition of SPD matrices).

(e) Using the same notation of the previous sections show that the solution of the regularized LS problem is given by

$$\hat{\mathbf{x}}=\sum_{i=1}^{\min(m,n)}\frac{\sigma_i}{\sigma_i^2+\lambda}\alpha_i\mathbf{v}_i.$$

(f) Read the deblurring example shown in the tutorials (Section 5.9.1) and try to explain it using the results above. Assume that the noise in the image corresponds to a singular vector $\mathbf{u}_i$ with a small corresponding singular value σ_i.

5. **Camera Calibration**

Given 3D world coordinates (x,y,z), their projection according to the pinhole camera model on an image plane (u,v) is:

$$\begin{bmatrix}u\\v\end{bmatrix}=\boldsymbol{K}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

Where

$$K=\begin{bmatrix}f_x&0&x_0\\0&f_y&y_0\\0&0&1\end{bmatrix}$$

The parameters f_x,f_y,x_0,y_0 are called the *intrinsic parameters* of the camera.
Particularly, f_x and f_y denote the focal length of the camera, and x_0,y_0 are the offset of the principal point relative to the origin.
For more information regarding the *pinhole camera* model, please refer to:

`https://en.wikipedia.org/wiki/Pinhole_camera_model`

You are given n correspondences, i.e., for every $i\in\{1,...,n\}$, you are given the pair:

$$\left(\begin{bmatrix}u_i\\v_i\end{bmatrix},\begin{bmatrix}x_i\\y_i\\z_i\end{bmatrix}\right)$$

(a) What is the minimal number of correspondences required to find K?

(b) Assuming the number of correspondences, n, is larger than the minimum number you found in the previous section, how would you solve the problem to obtain a good solution for K?
Show all your computations and derivations for the optimal solution in a least squares sense.

Homework submission instructions

1. The work is done in pairs. The course is large and there won't be exceptions.
2. **All the computations can be done by code. No need for manual ones.**
3. **All the code used for preparing the work should be documented and submitted as part of the work. That said, all the answers to the questions should be written in the text (in words), and not as part of the code or running scripts. We won't search for the answers in your scripts or printouts.**
4. **All graphs should include titles, legends, and axis labels.**
5. It is preferable to use either Julia, MATLAB, or Python, but any other language is OK.
6. It is preferable that the HW is prepared electronically, but scanned hand-written works will also be accepted as long as they are readable.
7. The work should be submitted in a ZIP file through the CS submission system.
8. Late submissions without penalty will be approved in cases of reserve duties or illness.
9. In other cases, the penalty of a late submission is 2^n where n is the number of days beyond the due date. That is 2 points for one day late, 4 for two, 8 for three and so on. (Fri-Sat counts as one day). **Such late submissions do not require approval.** For now we will not enforce this rule due to the rough situation. Try not to take too much advantage of that.