Problem_set_1

Jan, 2021

Instructions: Hand-in is in electronic form via email. Please create a single PDF-file containing all your answers and results. Show the names and student numbers of the group members on top. Maximum 4 students can work in a group. Make use of figures and tables. The code should be well documented and readable. Have a look at Latex-Example_with_Matlab (on Olat) for how to include Matlab code in a Latex file. Pay attention to the use of mcode package.

Exercise 1: A Linear Equation System

The simple example from the lecture was $\mathbf{Ax} = \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 4 & 1 \\ 1 & -1 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 30 \\ 40 \\ 15 \end{bmatrix}$$

Solve the system with the following methods

a) MATLAB's backslash-operator, \. Is this a direct or indirect solution method?

b) Compute the solution using the Gauss-Seidel iteration algorithm. Use the version explained in the lecture notes.

c) Compute the condition number, $\kappa(\mathbf{A})$, using `cond`. Interpret it. Is it high?

Exercise 2: Loops

a) Recall that machine epsilon, ϵ_m, is the smallest number for a given architecture such that $\epsilon_m + 1 > 1$. Develop an algorithm using a `while` loop that determines ϵ_m on your machine. Compare to `eps`. Do the same using a `for` loop.

b) Create two random matrices, $\mathbf{A}$, $\mathbf{B}$, each of size 400×600 (use `rand`). Compute $\mathbf{C} = \mathbf{A} + \mathbf{B}$ and $\mathbf{D} = \mathbf{A}'\mathbf{B}$ using a Matlab's direct implementation. In addition, programm a `for` loop that performs the same computation. How do both methods compare in speed (hint: `tic,toc`)?

Exercise 3: Vectors and random numbers

a) Fill $\mathbf{x}(n, 1)$, $n = 10$, with random numbers from a standard normal distribution (`randn`).

b) Calculate the mean and the standard deviation of $\mathbf{x}$.

c) Create a random variable $\mathbf{z}(n, 1) \sim \mathcal{N}(1, 2)$. Recall that this can be done by $z = x\sigma + \mu$.

d) Calculate the mean and the standard deviation of $\mathbf{z}$.

e) Use a `for` loop to redo steps c) and d) for $n = 1{,}000$ and for $n = 1{,}000{,}000$. Compare. What kind of error do we get for small n?

Exercise 4: Ordinary least squares (OLS) regression

Recall the linear regression model $\mathbf{y} = \mathbf{X}\beta + \epsilon$, where $\mathbf{y}$ is the $n \times 1$ vector of dependent variables, $\mathbf{X}$ is the $n \times p$ matrix of regressors, β is the $p \times 1$ vector of regression coefficients (with β_1 the intercept), and ϵ is the $n \times 1$ vector of errors. The ordinary least squares (OLS) estimator is $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. In this exercise, we will create our own artificial observations, for which we know the true β, and use OLS regression to estimate $\hat{\beta}$.

a) Let the true $\beta = [2, 3, 5]'$, and let $n = 10$. Create a random vector $\epsilon \sim \mathcal{N}(0, 1)$. Create $\mathbf{X} = [\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2]$ where $\mathbf{x}_1, \mathbf{x}_2$ are uniformly distributed random variables on the interval $[0, 100]$ (`rand`). With all these, create the dependent variables $\mathbf{y}$.

b) Compute $\hat{\beta}$ as shown above.

c) Compute β^* by solving $\mathbf{y} = \mathbf{X}\beta$ with the backslash operator. Compare.

d) Read `doc mldivide`.

e) Redo steps a) and c) for $n = 100000$.

Exercise 5: Filter Data

In this exercise you will work with real-world, time-series data on worker flows, apply the Hodrick-Prescott (HP) filter, and visualize the data. The HP filter is a tool to separate the cyclical component of a time series from its long-run trend.

a) Load the German worker flow data (file: *workerflows.xls*) provided by Jung and Kuhn (2014): Labor Market Institutions and Worker Flows: Comparing Germany and the U.S, in: *Economic Journal*.

b) Search the Internet for a Matlab file that implements the Hodrick-Prescott (HP) filter. Try to find a trustworthy source, download it, and read the usage description. Use a smoothing parameter of $\lambda = 14400$ and decompose the log of the employment to unemployment (EU) transition rate and the log of the unemployment to employment (UE) transition rate into a cyclical and a trend component.

c) Visualize the data with the following three figures: (i) EU raw data and corresponding HP Trend, (ii) UE raw data and corresponding HP trend, and (iii) cyclical component of EU and UE transition rates. Use the documentation to change plotting parameters (`doc plot`), including line width, line color, and line style.

Git

Git is a version control system. The benefits of using Git (or other version control systems):

- Easy way of tracking chronological versions
- Allows to develop different versions parallel (branching)
- Easy collaboration
- Also a way to backup work