

18.335 Problem Set 2

Due March 17, 2023 at 11:59pm. You should submit your problem set **electronically** on the 18.335 Gradescope page. Submit **both** a *scan* of any handwritten solutions (I recommend an app like TinyScanner or similar to create a good-quality black-and-white “thresholded” scan) and **also** a *PDF printout* of the Julia notebook of your computer solutions. A **template Julia notebook is posted** in the 18.335 web site to help you get started.

Problem 0: Pset Honor Code

Include the following statement in your solutions:

I will not look at 18.335 pset solutions from previous semesters. I may discuss problems with my classmates or others, but I will write up my solutions on my own. <your signature>

Problem 1: Stability and conditioning for linear systems

- (a) (From Trefethen and Bau, Exercise 18.1.) Consider the example

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \\ 1 & 1.0001 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 0.0001 \\ 4.0001 \end{pmatrix}.$$

- What are the matrices A^+ and P in this example? Give exact answers.
 - Find the exact solutions x and $y = Ax$ to the least squares problem $x = \arg\min_v \|Av - b\|_2$.
 - What are $\kappa(A)$, θ , and η ? Numerical answers computed with, e.g., Julia, are acceptable.
 - What numerical values do the four condition numbers of Theorem 18.1 take for this problem?
 - Give examples of perturbations δb and δA that approximately attain these four condition numbers.
- (b) (From Trefethen and Bau, Exercise 21.6.) Suppose $A \in \mathbb{C}^{m \times m}$ is *strictly column diagonally dominant*, which means that for each

column index k ,

$$|a_{kk}| > \sum_{j \neq k} |a_{jk}|.$$

Show that if Gaussian elimination with partial pivoting is applied to A , no row interchanges take place.

- (c) (From Trefethen and Bau, Exercise 23.2.) Using the proof of Theorem 16.2 as a guide, derive Theorem 23.3 from Theorems 23.2 and 17.1. In other words, show that solving symmetric positive definite (SPD) linear systems with a Cholesky factorization followed by forward and backward substitution is backward stable.

Problem 2: Banded factorization of a finite-difference matrix

Consider the advection equation $\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0$ with $(t, x) \in [0, T] \times [0, 1]$ for some $T > 0$. Let $u(x, t)$ have initial condition $u(x, 0) = b(x)$ and boundary conditions $u(0, t) = u(1, t) = 0$. For numerical approximation with finite-differences at equispaced times $0 < t_1, t_2, \dots, t_m < T$ and points $0 < x_1, x_2, \dots, x_n < 1$, we can approximate

$$\frac{\partial u_{jk}}{\partial t} \approx \frac{u_{jk} - u_{j(k-1)}}{\Delta t},$$

where $u_{jk} = u(x_j, t_k)$, $\Delta t = t_k - t_{k-1}$, $\Delta x = x_j - x_{j-1}$, and

$$\frac{\partial u_{jk}}{\partial x} \approx \frac{-u_{(j+2)k} + 8u_{(j+1)k} - 8u_{(j-1)k} + u_{(j-2)k}}{12\Delta x}.$$

- (a) Show that the vector $u_k = (u(t_k, x_1), u(t_k, x_2), \dots, u(t_k, x_n))^T$, representing an approximate solution on the grid at time t_k , can be obtained from u_{k-1} by solving the $n \times n$ linear system $(I + \sigma D)u_k = u_{k-1}$, where $\sigma = \Delta t / \Delta x < 1$ and

$$D = \begin{pmatrix} 0 & 2/3 & 1/12 & & \\ -2/3 & 0 & 2/3 & \ddots & \\ -1/12 & -2/3 & 0 & \ddots & 1/12 \\ & \ddots & \ddots & \ddots & 2/3 \\ & & -1/12 & -2/3 & 0 \end{pmatrix}.$$

- (b) Describe a banded elimination algorithm to factor the matrix $A = I + \sigma D$ into the product $A = LDU$: here, L and U^T are lower triangular with unit diagonal and have the same lower bandwidth as A , while D is a diagonal matrix. How does the memory cost (number of nonzeros stored explicitly) and flop count for your algorithm scale as $n \rightarrow \infty$? Implement your algorithm in the accompanying Julia notebook.
- (c) Can your algorithm fail without pivoting when $\sigma \leq 1$? How should one refine the approximation in space (decrease Δx) and in time (decrease Δt) to maintain stability without row exchanges? Explain why. (Hint: Problem 1(b) may be useful here.)
- (d) Use $A = LDU$ to compute solutions u_1, u_2, \dots, u_k , with the problem parameters set in the accompanying Julia notebook.

Problem 3: Regularized least-squares

Consider the regularized least-squares problem with regularization parameter $\lambda > 0$, given by

$$x_* = \operatorname{argmin}_x \|Ax - b\|_2^2 + \lambda \|x\|_2^2.$$

- (a) Show that the regularized least-squares problem is equivalent to a standard least-squares problem,

$$x_* = \operatorname{argmin}_x \left\| \begin{pmatrix} A \\ \sqrt{\lambda} I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \right\|_2^2.$$

- (b) If the SVD of A is $A = U\Sigma V^*$, show that the unique solution is

$$x_* = V(\Sigma^* \Sigma + \lambda I)^{-1} \Sigma^* U^* b.$$

- (c) Under what conditions on A does the regularized solution converge to the usual least-squares solution $\operatorname{argmin}_x \|Ax - b\|_2^2$ in the limit $\lambda \rightarrow 0$?
- (d) Describe a structure-exploiting Givens-based QR solver for the equivalent standard least-squares problem in part (a). How does the flop count compare to the standard QR solver for $\operatorname{argmin}_x \|Ax - b\|_2^2$?
- (e) (Not for credit.) Play with the example in the accompanying Julia notebook!