18.1. Consider the example

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

(a) What are the matrices A^+ and P for this example? Give exact answers.
(b) Find the exact solutions x and $y = Ax$ to the least squares problem $Ax \approx b$.
(c) What are $\kappa(A)$, θ, and η? From here on, numerical answers are acceptable.
(d) What are the four condition numbers of Theorem 18.1?
(e) Give examples of perturbations δb and δA that approximately attain these four condition numbers.

(a)

$$\begin{aligned} &\text{Given } A \in \mathbb{C}^{m\times n} \text{ of full rank, } m \ge n,\ b \in \mathbb{C}^m, \\ &\text{find } x \in \mathbb{C}^n \text{ such that } \|b - Ax\| \text{ is minimized.} \end{aligned} \tag{18.1}$$

The solution x and the corresponding point $y = Ax$ that is closest to b in range(A) are given by

$$x = A^+b, \qquad y = Pb, \tag{18.2}$$

where $A^+ \in \mathbb{C}^{n\times m}$ is the pseudoinverse (11.11) of A and $P = AA^+ \in \mathbb{C}^{m\times m}$ is the orthogonal projector onto range(A). We consider the conditioning of

$$A^+ = (A^*A)^{-1}A^* \in \mathbb{C}^{n,m}. \tag{11.11}$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1+10^{-4} \\ 1 & 1+10^{-4} \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1+10^{-4} & 1+10^{-4} \end{pmatrix}$$

$$\Rightarrow A^*A = A^TA = \begin{pmatrix} 3 & 3+2\cdot 10^{-4} \\ 3+2\cdot10^{-4} & 3+4\cdot10^{-4}+2\cdot10^{-8} \end{pmatrix}$$

$$\Rightarrow (A^*A)^{-1} = \underbrace{5\,0\cdot 10^{6}}_{\det^{-1}} \begin{pmatrix} 3+4\cdot10^{-4}+2\cdot10^{-8} & -3-2\cdot10^{-4} \\ -3-2\cdot10^{-4} & 3 \end{pmatrix}$$

$$\Rightarrow (A^* A)^{-1} = \begin{pmatrix} 150020001 & -150010000 \\ -150010000 & 15 \cdot 10^7 \end{pmatrix}$$

$$\Rightarrow A^+ = (A^* A)^{-1} A^* = \begin{pmatrix} 150020001 & -150010000 \\ -150010000 & 15 \cdot 10^7 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1+10^{-4} & 1+10^{-4} \end{pmatrix}$$

$$= \begin{pmatrix} 10.001 & -5 & -5 \\ -10 & 5 & 5 \end{pmatrix} \cdot 10^3$$

$$P = AA^+ = \begin{pmatrix} 1 & 1 \\ 1 & 1+10^{-4} \\ 1 & 1+10^{-4} \end{pmatrix} \cdot \begin{pmatrix} 10.001 & -5 & -5 \\ -10 & 5 & 5 \end{pmatrix} \cdot 10^3$$

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \frac{1}{2}$$

(b) $x = A^+ b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $y = Pb = \begin{pmatrix} 2 \\ 2+10^{-4} \\ 2+10^{-4} \end{pmatrix} = Ax$

(c) Using Numpy here, code in submission.

$\kappa(A) \approx 42429.235416083044$

$\theta \approx 0.684702873261185$

$\eta \approx 1 + 8.37278 \cdot 10^{-10}$

(d)

	y	x
b	$\dfrac{1}{\cos\theta}$	$\dfrac{\kappa(A)}{\eta\cos\theta}$
A	$\dfrac{\kappa(A)}{\cos\theta}$	$\kappa(A) + \dfrac{\kappa(A)^2\tan\theta}{\eta}$

here:

	y	x
b	1.2709772360 78942	54775.1770207547
A	54775.17706639765	1469883252.449082

(e)

$$\kappa_{b \mapsto y} = \frac{\|P\|}{\|y\|/\|b\|} = \frac{1}{\cos\theta}.$$

This establishes the upper-left result of Theorem 18.1. The condition number is realized (that is, the supremum in (12.5) is attained) for perturbations δb with $\|P(\delta b)\| = \|\delta b\|$, which occurs when δb is zero except in the first n entries.

Example: $\delta b = \sum_{i=1}^{n} \vec{e}_i$

$$\kappa_{b \mapsto x} = \frac{\|A^+\|}{\|x\|/\|b\|} = \|A^+\| \frac{\|b\|}{\|y\|} \frac{\|y\|}{\|x\|} = \|A^+\| \frac{1}{\cos\theta} \frac{\|A\|}{\eta} = \frac{\kappa(A)}{\eta\cos\theta}.$$

This establishes the upper-right result of Theorem 18.1. Here, the condition number is realized by perturbations δb satisfying $\|A^+(\delta b)\| = \|A^+\|\,\|\delta b\| = \|\delta b\|/\sigma_n$, which occurs when δb is zero except in the nth entry (or perhaps also in other entries, if A has more than one singular value equal to σ_n).

Example: $\delta b = \vec{e}_n$

For the lower two: Example: $\delta A = \begin{pmatrix} 0 & 0 \\ 0 & -10^{-4} \\ 0 & -10^{-4} \end{pmatrix}$

which makes A ultimately singular (rank 1).

19.1. Given $A \in \mathbb{C}^{m\times n}$ of rank n and $b \in \mathbb{C}^m$, consider the block 2×2 system of equations

$$\begin{bmatrix} I & A \\ A^* & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}, \qquad (19.4)$$

where I is the $m \times m$ identity. Show that this system has a unique solution $(r, x)^T$, and that the vectors r and x are the residual and the solution of the least squares problem (18.1).

Given $A \in \mathbb{C}^{m\times n}$ of full rank, $m \geq n$, $b \in \mathbb{C}^m$,
find $x \in \mathbb{C}^n$ such that $\|b - Ax\|$ is minimized. (18.1)

The solution x and the corresponding point $y = Ax$ that is closest to b in range(A) are given by

$$x = A^+ b, \qquad y = Pb, \qquad (18.2)$$

It is immediately obvious that $(r, x)^T$ is the unique solution, because

$$M = \begin{pmatrix} I & A \\ A^* & 0 \end{pmatrix} \text{ is non-singular.}$$

This follows directly from A being of full rank.

$$M \begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad Ir + Ax = b \iff r = b - Ax$$

Which is the definition of the residuals.

And $\quad A^* r = 0$ ← Those are the normal equations!

$$\Rightarrow \quad A^* r = A^* b - A^* A x = 0$$

$$\iff x = (A^* A)^{-1} A^* b = A^+ b$$

$\iff$ x is the solution! $\square$

20.3. Suppose an $m \times m$ matrix A is written in the block form $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where A_{11} is $n \times n$ and A_{22} is $(m-n) \times (m-n)$.
Assume that A satisfies the condition of Exercise 20.1.
(a) Verify the formula

$$\left[\begin{array}{c|c} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{array}\right] \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{array}\right]$$

for "elimination" of the block A_{21}. The matrix $A_{22} - A_{21}A_{11}^{-1}A_{12}$ is known as the *Schur complement* of A_{11} in A.
(b) Suppose A_{21} is eliminated row by row by means of n steps of Gaussian elimination. Show that the bottom-right $(m-n) \times (m-n)$ block of the result is again $A_{22} - A_{21}A_{11}^{-1}A_{12}$.

20.1. Let $A \in \mathbb{C}^{m \times m}$ be nonsingular. ~~Show that~~ A has an LU factorization if and only if for each k with $1 \le k \le m$, the upper-left $k \times k$ block $A_{1:k,1:k}$ is nonsingular. (Hint: The row operations of Gaussian elimination leave the determinants $\det(A_{1:k,1:k})$ unchanged.) Prove that this LU factorization is unique.

(a)

There's not a lot to verify. The upper-triangular entries are a direct result of matrix multiplication, the lower left entries are:

$$\left[\begin{array}{c|c} I & 0 \\ \hline \boxed{-A_{21}A_{11}^{-1} \quad I} \end{array}\right] \left[\begin{array}{cc} \boxed{\begin{matrix} A_{11} \\ A_{21} \end{matrix}} & \begin{matrix} A_{12} \\ A_{22} \end{matrix} \end{array}\right] = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \boxed{0} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{array}\right]$$

$$\left(-A_{21}A_{11}^{-1}, I\right) \cdot \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix} = -A_{21}\underbrace{A_{11}^{-1}A_{11}}_{I} + I\,A_{21}$$

$$= -A_{21} + A_{21} = \boxed{0}$$

(b) Well, that's exactly what the formula IS: The first n steps of Gaussian Elimination.

$A_k = L_{k-1} L_{k-2} \cdots L_1 A$

In each step we multiply A_k by L_k from the left:

$$x_k = \begin{bmatrix} x_{1k} \\ \vdots \\ x_{kk} \\ x_{k+1,k} \\ \vdots \\ x_{mk} \end{bmatrix} \xrightarrow{L_k} L_k x_k = \begin{bmatrix} x_{1k} \\ \vdots \\ x_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

To do this we wish to subtract ℓ_{jk} times row k from row j, where ℓ_{jk} is the *multiplier*

$$\ell_{jk} = \frac{x_{jk}}{x_{kk}} \qquad (k < j \le m). \qquad (20.6)$$

$\Rightarrow l_k = \frac{\tilde{x}_k}{x_{kk}}$

The matrix L_k takes the form

$$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\ell_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\ell_{mk} & & & 1 \end{bmatrix},$$

$\tilde{x}_k = (0, 0, \cdots, x_{kk}, x_{k+1,k}, \cdots)^T$

$$L = L_1^{-1} L_2^{-1} \cdots L_{m-1}^{-1} = \begin{bmatrix} 1 & & & & \\ \ell_{21} & 1 & & & \\ \ell_{31} & \ell_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{m1} & \ell_{m2} & \cdots & \ell_{m,m-1} & 1 \end{bmatrix}.$$

So $\underbrace{L_{m-1} L_{m-2} \cdots L_1}_{L^{-1}} A = U$

The first n steps: $\underbrace{L_n L_{n-1} \cdots L_1}_{=: \tilde{L}_n} A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & S \end{pmatrix}$

We know that $\tilde{L}_n$ has the form

$$\begin{pmatrix} I & 0 \\ B & I \end{pmatrix}$$

and to achieve the 0 in the lower left:

$$B A_{11} + I A_{21} \overset{!}{=} 0 \Rightarrow B = -A_{21} A_{11}^{-1}$$

$$\Rightarrow S = B A_{12} + I A_{22} = -A_{21} A_{11}^{-1} A_{12} + A_{22} \qquad \square$$

20.4. Like most of the algorithms in this book, Gaussian elimination involves a triply nested loop. In Algorithm 20.1, there are two explicit **for** loops, and the third loop is implicit in the vectors $u_{j,k:m}$ and $u_{k,k:m}$. Rewrite this algorithm with just one explicit **for** loop indexed by k. Inside this loop, U will be updated at each step by a certain rank-one outer product. This "outer product" form of Gaussian elimination may be a better starting point than Algorithm 20.1 if one wants to optimize computer performance.

Algorithm 20.1. Gaussian Elimination without Pivoting

$U = A, \quad L = I$
for $k = 1$ **to** $m - 1$
    for $j = k + 1$ **to** m
        $\ell_{jk} = u_{jk}/u_{kk}$
        $u_{j,k:m} = u_{j,k:m} - \ell_{jk} u_{k,k:m}$

I implemented this as part of my solution for Problem 6. (See my method lu_factorize_inplace in solve.py)

The pseudo-code looks like this:

for $k = 1$ to $m-1$

$$A_{k+1:m,k} = A_{k+1:m,k} / A_{k,k}$$

$$A_{k+1:m,k+1:m} = A_{k+1:m,k+1:m} - A_{k+1:m,k} \cdot A_{k,k+1:m}$$

23.1. Let A be a nonsingular square matrix and let $A = QR$ and $A^*A = U^*U$ be QR and Cholesky factorizations, respectively, with the usual normalizations $r_{jj}, u_{jj} > 0$. Is it true or false that $R = U$?

With these normalizations: Yes, it is true! and U being upper-triangular

Proof:

$$A^*A = R^*\underbrace{Q^*Q}_{I}R = R^*R = U^*U$$

$\Rightarrow$ if $r_{jj}, u_{jj} > 0$ and U is chosen to be upper-triangular, then $U = R$ $\square$

6.

positive definite!

$$A = R^* R$$

$$Ax = b \Rightarrow R^* \underbrace{R x}_{y} = b$$

$$y = Rx$$