MATH 110: LINEAR ALGEBRA
HOMEWORK #8

DAVID ZYWINA

§First set of problems

Problem 6. *If P is a permutation matrix, and $Y = XP$, show that Y has the same columns as X but in a permuted order.*

For each $j \in \{1, \cdots, n\}$, let $\sigma(j)$ be the unique element of $\{1, \cdots, n\}$ such that $P_{\sigma(j),j} = 1$ (i.e. the unique 1 in the jth column of X occurs in the $\sigma(j)th$ row). Since P is a permutation matrix, the numbers $\sigma(1), \cdots, \sigma(n)$ are a permutation of the numbers $1, \cdots, n$.

Fix a $j \in \{1, \cdots, n\}$. For each i,

$$(XP)_{i,j} = \sum_{k=1}^{n} X_{i,k} P_{k,j} = X_{i,\sigma(j)} P_{\sigma(j),j} = X_{i,\sigma(j)}.$$

This shows that the jth column of XP is the $\sigma(j)$th column of X. Therefore Y has the same columns as X but in a permuted order.

Problem 7. *Show that if P_1 and P_2 are permutation matrices, then so is P_1P_2.*

The previous problem (with $X = P_1$ and $P = P_2$) shows that P_1P_2 has the same columns as P_1 but in a permuted order. Since P_1 has the same columns as the identity matrix I (possibly permuted), this shows that the columns of P_1P_2 are just a permutation of those of I. Thus P_1P_2 is a permutation matrix.

Problem 8. *Show that if P is a permutation matrix, so is P^t, and $P^t = P^{-1}$.*

A *permutation matrix*, by definition, is an $n \times n$ matrix with exactly one 1 in each row, one 1 in each column, and the other entries equal to 0. Since P^t is the matrix where the rows are swapped with columns, and columns are swapped with rows it is immediate that P^t must also be a permutation matrix.

We can write $P = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$, where $\{c_1, \cdots, c_n\} = \{e_1, \cdots, e_n\}$ (e_i is the ith column of I). The transpose of P can be expressed as

$$P^t = \begin{bmatrix} c_1^t \\ c_2^t \\ \vdots \\ c_n^t \end{bmatrix}.$$

Date: October 27th.

For $1 \le i, j \le n$,

$$(P^t P)_{i,j} = c_i^t c_j = \delta_{i,j}$$

(the last equality follows from the observation that for $1 \le i, j \le n$, $e_i^t e_j = \delta_{i,j}$). So $P^t P = I$ and hence P is invertible with inverse P^t.

Problem 9. *Use LU decomposition to compute an LU decomposition of*

$$A = \begin{bmatrix} 0 & 2 & 4 \\ 0 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}$$

and use it to describe the complete solution set of $Ax = \begin{bmatrix} 8 \\ 4 \\ -4 \end{bmatrix}$ *and* $Ax = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

We need to have a non-zero element in the top left corner of our matrix to start, so we multiply A by the permutation matrix $P = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ which interchanges the first and third row; we get

$$PA = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix}.$$

The first column of PA is already done, so we now work on the second column. Subtracting 2 times the second row from the third row, gives us the matrix:

$$U' = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

We let

$$L' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

(where the 2 is there because we multiplied the 2nd row by 2 before subtracting it from the 3rd row).

You can now check that we indeed have $PA = L'U'$ as desired. So multiplying both sides by P^{-1} $(= P^t = P)$, we have

$$A = PL'U' = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

You might argue (rightly) that this is not the LU decomposition defined in class. Since A has rank 2, L should be a 3×2 matrix and U should be a 2×3 matrix. This can be achieved by simply removing the last column of L' and the last row of U' (This might seem slightly tricky but it can always be done in this situation. Convince yourself that when multiplying the matrices the removed column and row contribute nothing.).

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = PLU.$$

Now that we have computed the LU decomposition we can solve $Ax = PLUx = \begin{bmatrix} 8 \\ 4 \\ -4 \end{bmatrix}$. We first solve the equation (where $y = Ux$)

$$Ly = P^{-1} \begin{bmatrix} 8 \\ 4 \\ -4 \end{bmatrix} = P \begin{bmatrix} 8 \\ 4 \\ -4 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \\ 8 \end{bmatrix}.$$

This is the same as the equations:

$$\begin{aligned} y_1 &= -4 \\ y_2 &= 4 \\ 2y_2 &= 8. \end{aligned}$$

These equations are consistent and have solution $y = \begin{bmatrix} -4 \\ 4 \end{bmatrix}$. It remains to solve the equation $Ux = y = \begin{bmatrix} -4 \\ 4 \end{bmatrix}$. This is the same as the equations:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= -4 \\ x_2 + 2x_3 &= 4. \end{aligned}$$

The second equation shows that $x_2 = 4 - 2x_3$. Substituting into the first equation we get $3x_1 = -4 - 2(4 - 2x_3) - x_3 = -12 + 3x_3$, thus $x_1 = -4 + x_3$. Therefore the solutions to $Ax = \begin{bmatrix} 8 \\ 4 \\ -4 \end{bmatrix}$ are

$$\left\{ \begin{bmatrix} -4 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

We finally solve $Ax = PLUx = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. We first solve the equation (where $y = Ux$)

$$Ly = P^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = P \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

This is the same as the equations:

$$\begin{aligned} y_1 &= 0 \\ y_2 &= 1 \\ 2y_2 &= 0. \end{aligned}$$

However, these equations are not consistent (they imply that $y_2 = 0$ and 1). Thus the original equation $Ax = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ has *no* solutions.

§Second set of problems

Problem 1. *Use "substitution" to solve $Lx = b$ where*

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix}.$$

Writing $x = [x_1, x_2, x_3]^t$ the equation $Lx = b$ becomes:

$$\begin{aligned} x_1 &= 3 \\ x_2 + 2x_1 &= -4 \\ x_3 + 3x_2 + 4x_1 &= 0. \end{aligned}$$

Solving these equations is easy. We have immediately that $x_1 = 3$. Substituting this into the second equation and solving for x_2 gives us, $x_2 = -10$. Substituting the values of x_1 and x_2 into the third equation gives us, $x_3 = 18$. Therefore the solution is

$$x = \begin{bmatrix} 3 \\ -10 \\ 18 \end{bmatrix}.$$

Problem 2. Use "substitution" to compute L^{-1}, where L is given above.

We can write $L^{-1} = [x\ y\ z]$, where x, y and z are column vectors. We then have

$$[e_1\ e_2\ e_3] = I = LL^{-1} = L[x\ y\ z] = [Lx\ Ly\ Lz].$$

Therefore we need only solve the equations $Lx = e_1$, $Ly = e_2$ and $Lz = e_3$.
We now solve the first equation, $Lx = e_1$, which can be rewritten as:

$$\begin{aligned} x_1 &= 1 \\ x_2 + 2x_1 &= 0 \\ x_3 + 3x_2 + 4x_1 &= 0. \end{aligned}$$

We have immediately that $x_1 = 1$. Substituting this into the second equation and solving for x_2 gives us, $x_2 = -2$. Substituting the values of x_1 and x_2 into the third equation gives us,

$x_3 = 2$. Therefore $x = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$. Similarily, we also compute $y = \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix}$ and $z = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.
Putting everything together we get:

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}.$$

Problem 3. *In lecture we showed that the number of arithmetic operations needed to compute an LU decomposition of an $n \times n$ matrix with rank n was $\frac{2}{3}n^3$ + lower order terms.*

(3a) *How many more arithmetic operations does it take, given the LU decomposition of A, to compute A^{-1}? The method to use is for each column e_j of the identity matrix, solve $Ax = e_j$ for x = column j of A^{-1}. Use substitution with the L and U factors to solve this equation. Recall that the P_L and P_R factors only involve reordering, no arithmetic operations. Your answer should be of the form "$c_1 n^{c_2}$ + lower order terms", where c_1 and c_2 are constants you need to determine. Add $\frac{2}{3}n^3$ to your answer to determine the total number of arithmetic operations to compute the inverse of an invertible $n \times n$ matrix.*

To compute A^{-1} we need to find column vectors $c_1, \cdots, c_n$ such that

$$A \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix} = I.$$

So computing $A^{-1} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$ is equivalent to solving the equations

$$Ac_1 = e_1, \cdots, Ac_n = e_n.$$

The next Lemma analyzes the cost of solving $Ax = b$ for a general b. We can multiply this cost by n to get the cost of solving the n linear systems $Ac_i = e_i$ for $i = 1$ to n. Then we show that by exploiting all the zeros in e_i, we can actually solve $Ac_i = e_i$ more cheaply.

Lemma 1. Suppose we have an LU-decomposition, $A = P_L L U P_R$, where A is an $n \times n$ invertible matrix.
Then it takes $2n^2 - n$ arithmetic operations to solve an equation

$$Ax = b.$$

Proof. The equation is the same as $LU(P_R x) = P_L^t b$. Since multiplying by a permuation matrix takes no operations (it involves only swapping entries), we may assume our equation is of the form

$$LUx = b.$$

We first solve the equation $Ly = b$ for $y = [y_1, \cdots, y_n]^t$. Since L is a lower triangular matrix with 1's on the diagonals, the equation can be rewritten as:

$$\begin{aligned} y_1 &= b_1 \\ y_2 + L_{2,1}y_1 &= b_2 \\ y_3 + L_{3,2}y_2 + L_{3,1}y_1 &= b_3 \\ y_4 + L_{4,3}y_3 + L_{4,2}y_2 + L_{4,1}y_1 &= b_4 \\ &\vdots \\ y_n + L_{n,n-1}y_{n-1} + \cdots + L_{n,2}y_2 + L_{n,1}y_1 &= b_n \end{aligned}$$

The first equation requires 0 operations to solve for y_1.
The second equation requires 2 operations to solve for y_2 (1 mult. and 1 subtractions).
The third equation requires 4 operations to solve for y_3 (2 mult. and 2 subtractions).
The fourth equation requires 6 operations to solve for y_4 (3 mult. and 3 subtractions).

$$\vdots$$

The final equation requires $2(n-1)$ operations to solve for y_n ($n-1$ mult. and $n-1$ subtractions).

Thus the number of arithmetic operations require to solve the equation $Ly = b$ is

$$0 + 2 + 4 + 6 + \cdots + 2(n-1) = 2(1 + 2 + 3 + \cdots + (n-1)) = 2\frac{(n-1)n}{2} = n^2 - n.$$

Now that we have solved for y, it remains to solve for x in the equation $Ux = y$. A similar process shows that solving $Ux = y$ takes n^2 operations. (The proof is the same except that U does not necessarily have 1's on the diagonal, so n extra divisions will be needed in general).

Putting everything together, it will take $(n^2 - n) + n^2 = 2n^2 - n$ arithmetic operations to solve $Ax = b$. $\square$

Using the approach, the cost of inverting A is the cost of LU decomposition ($\frac{2}{3}n^3$ + lower order terms) plus the cost of solving $Ac_1 = e_1$ through $Ac_n = e_n$ ($n \cdot (2n^2 - n) = 2n^3$ + lower order terms), or altogether $\frac{8}{3}n^3$ + lower order terms.

However, we can do better by taking advantage of the zeros in e_i. Since the first $i-1$ entries of e_i are zero, it is easy to see that the first $i-1$ entries of $y = L^{-1}e_i$ are zero (this is the same as saying that L^{-1} is also lower triangular). Therefore, we do not need to compute them. Thus when solving $Ly = e_i$ we can start with $y_i = 1$, then solve $L_{i+1,i}y_i + y_{i+1} = 0$ for $y_{i+1} = -L_{i+1,i}$, then solve $L_{i+2,i}y_i + L_{i+2,i+1}y_{i+1} + y_{i+2} = 0$ for y_{i+2}, and so on. This is the same as solving a triangular system of dimension $n-i+1$, for a cost of $(n-i+1)^2 + (n-i+1)$ by the same analysis as above. Summing this from $i = 1$ to n yields the cost of solving all the $Ly = e_i$: [1].

$$\sum_{i=1}^{n}[(n-i+1)^2+(n-i+1)] = \sum_{i=1}^{n}[i^2+i] = \frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2} = \frac{n^3}{3} + \text{ lower order terms}$$

[1]Recall: $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ and $1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$

The cost of the subsequent solves with U does not change (n^3) for a total inversion cost of $\frac{2}{3}n^3 + \frac{1}{3}n^3 + n^3 = 2n^3$ (plus lower order terms).
(3b) *Now suppose A is $m \times n$ with rank r (the general case). How many arithmetic operations are needed to compute the LU decomposition?*

Let $C(m, n, r)$ be the number of arithmetic operations required to compute the LU decomposition of an $m \times n$ matrix with rank r.

The following argument will be done using the inductive algorithm done in lecture. First of all, multiplying by permutation matrices requires no arithmetic operations (this just swaps rows and columns); so we can ignore the contributions from the permutation matrices since they do not affect the value of $C(m, n, r)$.

Let A be an $m \times n$ matrix of rank r. After a possible permuation, we may assume that

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix},$$

where $A_{1,1}$ is a *non-zero* scalar. We then define $X = A_{2,1}/A_{1,1}$ and $S = A_{2,2} - XA_{1,2}$. (For future reference: $A_{1,2}$ is $1 \times (n-1)$, $A_{2,1}$ is $(m-1) \times 1$, and $A_{2,2}$ is $(m-1) \times (n-1)$)

It takes $m - 1$ divisions to calculate X.
It takes $(m-1)(n-1)$ arithmetic operations (all multiplications) to compute $XA_{1,2}$, then $(m-1)(n-1)$ subtractions to compute S.
Note: So far we have done $(m-1) + 2(m-1)(n-1)$ arithmetic operations.

We then have

$$A = \begin{pmatrix} 1 & 0 \\ X & I_{m-1} \end{pmatrix} \begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & S \end{pmatrix}.$$

The matrix S is $(m-1) \times (n-1)$. Since A has rank r and $A_{1,1} \neq 0$ we see that S has rank $r - 1$. Thus it takes $C(m-1, n-1, r-1)$ arithmetic operations to compute an LU decomposition

$$S = P_{LS} L_S U_S P_{RS}.$$

In the lecture notes, it is then derived that

$$A = \begin{pmatrix} 1 & 0 \\ 0 & P_{LS} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ P_{LS}^t X & L_S \end{pmatrix} \begin{pmatrix} A_{1,1} & A_{1,2} P_{RS}^t \\ 0 & U_S \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_{RS} \end{pmatrix}.$$

This is the desired LU decomposition of A. Note that all of the multiplications in these matrices are by permutation matrices and, as explained earlier, take no arithmetic operations to perform.
Thus we have proven the recursive equation,

$$C(m, n, r) = (m-1) + 2(m-1)(n-1) + C(m-1, n-1, r-1).$$

The base case for our recursion is $C(m, n, 0) = 0$ (the zero matrix is the only $m \times n$ matrix with rank 0).

By repeated using the recursive relation until we get to our base case, we get the explicit formula:

$$\begin{aligned} C(m,n,r) = & \quad (m-1)+(m-2)+...+(m-r) \\ & + \quad 2(m-1)(n-1)+2(m-2)(n-2)+...+2(m-r)(n-r) \end{aligned}$$

We now write this in a more explicit form

$$\begin{aligned} C(m,n,r) &= \sum_{i=1}^{r}(m-i)+2\sum_{i=1}^{r}(m-i)(n-i) \\ &= rm-\sum_{i=1}^{r} i+2rmn-2(m+n)\sum_{i=1}^{r} i+2\sum_{i=1}^{r} i^2 \\ &= rm-\frac{r(r+1)}{2}+2rmn-2(m+n)\frac{r(r+1)}{2}+2\frac{r(r+1)(2r+1)}{6} \end{aligned}$$

To get a better idea on the growth of this value we collect all the higher order terms (the terms of degree 3).

$$C(m,n,r) = 2rmn-(m+n)r^2+\frac{2}{3}r^3+\text{ lower order terms }.$$

For the special case where $m=r=n$, we get

$$C(n) := C(n,n,n) = \frac{2}{3}n^3+\text{ lower order terms}$$

as in lecture.