18.330 Pset 6

Kathleen Brandes

April 4, 2019

1 Problem 1

The code for problem 1 is on stellar in problem1.jl.

1.1 Part A

This matrix is non-symmetric, so I used the standard LU decomposition to get:

$$L = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 2.4 & 1.0 & 0.0 \\ -1.0 & -0.285714 & 1.0 \end{bmatrix}$$

$$U = \begin{bmatrix} 1.0 & 3.0 & 2.0 \\ 0.0 & -10.5 & -3.7 \\ 0.0 & 0.0 & 2.94286 \end{bmatrix}$$

And the solution to the system of equations, $A\mathbf{x} = \mathbf{b}$ was:

$$x = \begin{bmatrix} -0.359223 & -0.427184 & 1.32039 \end{bmatrix}$$

1.2 Part B

Since this matrix is positive semidefinite and symmetric, then the cholesky decomposition will be faster. The resulting lower triangular matrix was:

$$L = \begin{bmatrix} 2.23607 & 0.0 & 0.0 \\ -1.34164 & 2.04939 & 0.0 \\ 0.894427 & -0.48795 & 1.99045 \end{bmatrix}$$

And the solution to the system of equations, $A\mathbf{x} = \mathbf{b}$ was:

$$x = \begin{bmatrix} 0.942883 & 2.07746 & 1.87702 \end{bmatrix}$$

1.3 Part C

This matrix has negative eigenvalues, and therefore is not positive semidefinite, so I used the standard LU decomposition to get the following:

$$L = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ -1.5 & 1.0 & 0.0 \\ 1.0 & -1.33333 & 1.0 \end{bmatrix}$$

$$U = \begin{bmatrix} 2.0 & -3.0 & 2.0 \\ 0.0 & -1.5 & 2.0 \\ 0.0 & 0.0 & 2.66667 \end{bmatrix}$$

And the solution to the system of equations, $A\mathbf{x} = \mathbf{b}$ was:

$$x = \begin{bmatrix} -0.5 & 1.0 & 2.5 \end{bmatrix}$$

2 Problem 2

2.1 Part A

To prove the inverse of an upper triangular matrix is also upper triangular, let $R \in R^{nxn}$ be upper triangular. Assuming an inverse exists, then let matrix $B = R^{-1}$ such that $BR = I$, where I is the identity matrix in R^{nxn}. We know that the i-th column of the inverse will be the solution to $R\mathbf{b}_i = \mathbf{e}_i$ where $\mathbf{e}_i$ will be the unit vector with 1 on the i-th element location and $\mathbf{b}_i$ is each column of the inverse matrix B such that $B = [b_1, b_2, ..., b_n]$. Then we can get a system of equations from the entire matrix operation of $BR = I$ and the above fact:

$$\begin{gathered} b_1 r_{11} = e_1 \\ b_1 r_{12} + b_2 r_{22} = e_2 \\ ... \\ b_1 r_{1n} + b_2 r_{2n} + ... + b_n r_{nn} = e_n \end{gathered}$$

Where this system can be simplified:

$$\begin{gathered} b_1 = \tfrac{e_1}{r_{11}} \\ b_2 = \tfrac{e_2 - b_1 r_{12}}{r_{22}} \\ ...\ b_n = \tfrac{e_n - \sum_{i=1}^{n-1} b_i r_{in}}{r_{nn}} \end{gathered}$$

In general, we have $b_j = \frac{e_j - \sum_{i=1}^{n-1} b_i r_{ij}}{r_{jj}}$. Therefore, since these vectors b_j have zeros for all components with index larger than j, we can see that $B = R^{-1}$ is also upper triangular.

2.2 Part B

Prove the product of two upper triangular matrices, R_1, R_2, is also upper triangular. We know that if $R = R_1R_2$ is the product, then entry $R_{ij} = r^1_{i1}r^2_{1j} + ... + r^1_{in}r^2_{nj}$. To have the product, R, be upper triangular, then all entries $r^1_{ik}r^2_{kj}$ where $i > j$ must be zero. If $i > k$, then r^1_{ik} will be zero since R_1 is upper triangular. If $k > j$, then r^2_{kj} will be zero since R_2 is also upper triangular. Therefore, when $i > j$, we have either $i > k$ or $k > j$ or both, so then all the entries for that given R_{ij} will be zero, making it upper triangular.

2.3 Part C

We have $A = Q_1R_1 = Q_2R_2$,, which we can multiply as $Q_2^T(Q_1R_1)R_1^{-1} = Q_2^T(Q_2R_2)R_1^{-1}$ to get $Q_2^TQ_1 = R_2R_1^{-1}$. Here $R_2R_1^{-1}$ is also upper triangular because both R_1 and R_2 are upper triangular, and we know from part a that R_1^{-1} is upper triangular and the product of two upper triangular matrices is also upper triangular as proven in part b.

Additionally, we can show that $A = Q_2^TQ_1$ is orthogonal because the product of an orthogonal matrix and its transpose is the identity ($B^TB = I$ for B orthogonal), therefore:

$$A^TA = (Q_2^TQ_1)^T(Q_2^TQ_1) = Q_1^T((Q_2^T)^TQ_2^T)Q_1 = Q_1^TIQ_1 = Q_1^TQ_1 = I$$

which is true because both Q_1 and Q_2 are orthogonal by definition.

Then, the only upper triangular matrix with positive diagonal entries that is also orthogonal is the identity matrix, which can show by taking successive inner products of the columns since they are orthonormal. If we consider the product of $Q_2^TQ_1$, let q_i^1 be the i-th column of Q_1 and similar notation for Q_2, then the result is successive inner products of matrices:

$$Q_2^TQ_1 = \begin{bmatrix} q_1^2 \cdot q_1^1 & q_1^2 \cdot q_2^1 & ... & q_1^2 \cdot q_n^1 \\ ... & & & \\ q_n^2 \cdot q_1^1 & q_n^2 \cdot q_2^1 & ... & q_n^2 \cdot q_n^1 \end{bmatrix}$$

Since this matrix must be positive and diagonal, then we know all entries $q_i^2 \cdot q_j^1$ where $i \neq j$ are zero, meaning these vectors are not orthogonal. And because $q_i^2 \cdot q_i^1$ is non-zero and positive, the two vectors both must be orthogonal to each other. therefore, we know that $q_i^2 \cdot q_i^1 = 1$ for all diagonal entries because the vectors are also orthonormal, so the inner product is 1. From this it follows that Q is unique for the QR decomposition.

2.4 Part D

Since in part c we show that $A = Q_2^TQ_1 = I$, then we also know that $R_2R_1^{-1} = I$. We can right multiply by R_1 to get:

$$R_2R_1^{-1}R_1 = IR_1 \; R_2I = IR_1 \; R_2 = R_1$$

Therefore, we can see that R must also be unique for the QR decomposition.

3 Problem 3

The code for problem 3 is on stellar in problem3.jl.

3.1 Part A

To get the eigenvectors of the matrix, we need to keep a running product of each iteration of $\hat{Q}_i$, so $\hat{Q}_k = \hat{Q}_{k-1} * ... * \hat{Q}_1 \hat{Q}_0$, because if the eig-qr function converges, then the matrix $\hat{Q}$ will converge to an orthonormal basis of eigenvectors. As well, we know that the eigenvalues can be extracted from the diagonal of the resulting A because it is triangular. Using this information, we can get the eigenvalues and respective eigenvectors for a matrix A using the eig-qr function.

3.2 Part B

Using the test matrix, we get eigenvalues $\{-5.6758, 4.91639, 2.75942\}$ and eigenvectors = $\{[0.412659, -0.560236, 0.718226]^T, [0.597434, 0.761669, 0.250865]^T, [-0.687594, 0.325571, 0.649013]^T\}$. These are checked as eigenvalues/vectors by comparing the results of Av and λv. Further, we can ensure that eigenvectors v_i are orthonormal by checking $v_i \cdot v_j = 1$ if $i = j$ and zero otherwise.

3.3 Part C

Using the second test matrix, we get eigenvalues $\{-5.38038, 5.07159, 2.30878\}$ and eigenvectors = $\{[0.344519, -0.636, 0.690514]^T, [0.609579, 0.710942, 0.350677]^T, [-0.713947, 0.300108, 0.632626]^T\}$. When checking if the eigenvectors and eigenvalues are valid for this matrix, we get that the second two eigenvalues $5.07159, 2.30878$(and respective vectors) are not correct for this matrix. These are likely incorrect because the eigenvectors for this matrix will be non-orthogonal since the matrix is nonsymmetric, and this algorithm will only find orthonormal basis for eigenvectors.

4 Problem 4

The code for problem 4 is on stellar in problem4.jl.

4.1 Part A

Eigenvalues and vectors for the original square matrix, A, can all be written as $Av = \lambda v$.We can transform this as such:

$$\begin{gathered}
Av = \lambda v \\
(A - \mu I)v = (\lambda - \mu)v \\
\tfrac{(A-\mu I)v}{(\lambda-\mu)} = v \\
\tfrac{v}{(\lambda-\mu)} = (A - \mu I)^{-1} v
\end{gathered}$$

We can subtract a constant from the matrix diagonal, and this will similar change the eigenvalues by that constant as well. Therefore, we get the desired result for inverse iteration.

4.2 Part B

Power iteration converges to the largest eigenvalue (and respective eigenvector) for the matrix A. Inverse iteration is simply power iteration for the matrix $(A - \mu I)^{-1}$. The eigenvalues for this matrix will be $\frac{1}{\lambda_i - \mu}$ where the λ_i's are the eigenvalues of the original matrix A. Therefore the new largest eigenvalue will be $max(\frac{1}{\lambda_i - \mu})$, which occurs when $\lambda_i \approx \mu$. Therefore, inverse iteration will converge to the eigenvalue closest to μ, provided that the initial guess b_0 has a component in the eigenspace, $\mu \neq \lambda_i, \forall i$, and the eigenvalue of A closest to μ is not repeated.

4.3 Part C

Using inverse iteration, we find the matrix has eigenvalues of $\{-5.38038, 2.30878, 5.07159\}$ and eigenvectors of $\{[-0.344519, -0.609579, 0.713947]^T, [0.772971, 0.127599, 0.621478]^T, [-0.62409, 0.730983, 0.275999]^T\}$. These results are checked in the same way as problem 3 (comparing the results of Av and λv and checking orthonormality with $v_i \cdot v_j = 1$ if $i = j$ and zero otherwise).