1 Problem 1

1.1 a

For a function $q(s) = a_m s^m + a_{m-1}s^{m-1}... + a_0$, we can say that $Q(A) = a_m A^m + a_{m-1}A^{m-1}... + a_0 I$. Since we know that the matrix A^k is equivalent to $V\Lambda^k V^{-1}$, this is equivalent to $Q(A) = a_n(V\Lambda^n V^{-1}) + a_{n-1}V\Lambda^{n-1}V^{-1}... + a_0 I$.

We know that for the same polynomial $q(s)$, we can say that when $q(\Lambda) = diag(q(\lambda_1)...q(\lambda_n)$, $Vq(\Lambda)V^{-1} = Vdiag(q(\lambda_1)...q(\lambda_n)V^{-1}$. This means that $Vq(\Lambda)V^{-1} = V(a_n\lambda_n + a_n\lambda_{n-1}... + a_0)V^{-1} + V(a_{n-1}\lambda_n + a_n\lambda_{n-2}... + a_0)V^{-1}... + V(a_0)V^{-1}$. This is directly equivalent to $q(A) = a_n V\Lambda^n V^{-1} + a_{n-1}V\Lambda^{n-1}V^{-1}... + a_0 I$, meaning the two are the same.

1.2 b

Since we know from a. that the two are the same, this means that $p(A) = Vp(\Lambda)V^{-1}$, where $p(\Lambda) = \lambda^n + a_{n-1}\lambda^{n-1}... + a_0$. This applies for every eigenvalue $\lambda_1...\lambda_n$, where this is equivalent to $V\Lambda^n V^{-1} + a_{n-1}V\Lambda^{n-1}V^{-1}... + a_0$, which we know is equal to 0.

2 Problem 2

2.1 a

If $A = A^T$, then if $A = VBV^{-1}$, $A^T = (VBV^{-1})^T = (V^{-1})^T B^T V^T$. Therefore, $A = VBV^{-1} = A^T = (V^{-1})^T B^T V^T$. If $V^T V = I$, then $(VBV^{-1})(V^{-1})^T = VB$, and $((V^{-1})^T B^T V^T)(V^{-1})^T = (VBV^{-1})^T (V^{-1})^T = (V^{-1}VBV^{-1})^T$. This does not lead us to the same equation as for A, meaning we cannot prove that $B = B^T$ from this, making it false.

2.2 b

Two matrices A and B are similar when there exists some matrix P such that $A = PBP^{-1}$ and $B = P^{-1}AP$. Then if $A = A^T$, $PBP^{-1} = (PBP^{-1})^T = (P^{-1})^T B^T P^T$. This must mean that $P = P^{-1}$ and $B = B^T$, meaning that B is therefore symmetric, so this is True.

2.3 c

This is true, because if a matrix is not diagonalizable, then that implies there are remaining eigenvalues that are equal 0 (the number of 0 eigenvalues is determined by rank-nullity). Summing products involving these eigenvalues that are equal to 0 will give us the same solution as before+0, or still 0. Therefore, this is true.

3 Problem 3

Let A be the following mxn matrix:

a_{11} a_{12} ... a_{1n}

a_{21} a_{22} ... a_{2n}

...a_{m1} a_{m2} ... a_{mn}

We know that for this matrix A, the max 1-norm, $||A||_1$, is equal to $||Ax||_1$ where $||x||_1 = 1$. Since x has dimension nx1, this means that we can write $||Ax||_1$ as $\sum_{i=1}^{m} ||(A_i)x|| <= \sum_{i=1}^{m} ||(A_i)|||x|$. We know that this is going to be maximized when the value of $|x| = 1$ because any value less than 1 will decrease the total sum, since they are decimal values less than 1.

We also know that this is maximized when we only include the maximum column i of A because that is the only column that will be multiplied by the vector with norm 1, x. Therefore, $||A||_1$ is equal to $||(A_i)||$ where A_i is the maximum column of A. This is identical to max $\sum_{i=1}^{m} |a_{ij}|$ for the maximum column of A j, or the max column sum.

4 Problem 4

For a stochastic matrix, the columns and rows (depending on column-stochastic and row-stochastic) will each add to 1 with no negative values. If a matrix is column-stochastic, then its columns, specifically, will add up to 1.

For column-stochastic matrix A, let us say we know that it has a set of eigenvalues λ_1... λ_n. We can say that there is some eigenvector, x, corresponding to the maximum value eigenvalue, λ so that $Ax = \lambda x$. For the product Ax, we know that since no element is less than 0, then the highest possible value of Ax will be equal to 1 times the maximal value of x. For the eigenvalue λ, the maximal value will be equal to λ times the maximal value of x. Therefore, λ has a maximum value of 1, otherwise it would be impossible for A to be a column=stochastic matrix.

5 Problem 5

If we say that $\hat{X}_k = U\Sigma V^T$ and $X_k = YZ^T$, then we can say V^T has columns that are equal to the eigenvectors of X^TX, by the definition of singular value decomposition. Similarly, we can say that the columns of U are equal to the eigenvectors of XX^T.

Since the matrix Σ is a diagonal matrix filled with values $\sigma_1...\sigma_k$ for a decomposition of rank k, we can say that $U\Sigma$ is equal to the eigenvectors of XX^T each scaled by a corresponding singular value (e.g. eigenvector one is scaled by σ_1... eigenvector k is scaled by σ_k).

We can say that this scaling is equivalent to projecting each column of U onto the space generated by eigenvalue k for a space of rank k, thereby creating a matrix with rank k. This matrix is equivalent to Y, and since V^T and Z^T have the same definition, we can say that the two are equivalent, so $\hat{X}_k = X_k$.

6 Problem 6

Attached to bottom of page in separate PDF.