HW1

1 - Least Square Solution

Find the best solution, in the least-squared sense, to the following system of equations:

$x1 - 2x2 = -2$
$x1 - 2x2 = 5$
$-2x1 + x2 = 1$
$x1 - 3x2 = -3$

We can convert the given system into a matrix equation:

$$\begin{bmatrix} 1 & -2 \\ 1 & -2 \\ -2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 1 \\ -3 \end{bmatrix}$$

Which takes the form $Ax = b$. Given the fact that this system is overdetermined, we must then find an optimal solution to the problem in a least squares sense. To find a least square solution to this problem we solve the equation $A^TAx = A^Tb$, for the system given above. This results in the following system of equations:

$$A^TAx = A^Tb$$
$$\begin{bmatrix} 1 & 1 & -2 & 1 \\ -2 & -2 & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -2 \\ -2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 & 1 \\ -2 & -2 & 1 & -3 \end{bmatrix} \begin{bmatrix} -2 \\ 5 \\ 1 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} 7 & -9 \\ -9 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Solving this system reveals a least squares solution to this problem. We find that a vector

$$x_1 = 0$$
$$x_2 = 0.222$$

provides the least square solution for our original system.

2 - Pseudo Inverse

If the unweighted pseudo inverse solution

$x = A^\dagger y = A^T(AA^T)^{-1}y$

minimizes the value of $\frac{1}{2}x^Tx$ subject to the constraint $Ax = y$, what does the weighted pseudo inverse solution

$x = A_B^\dagger y = BA^T(ABA^T)^{-1}y$

minimize under the same constraint?

Starting from the Hamiltonian definition:

$$H = \frac{1}{2}x^Tx + \lambda^T(Ax - y)$$

We can reconstruct exactly what function is being minimized by following the process for determining the pseudo inverse. In our case, all the final A^T matrices are premultiplied by a weighting matrix B. This weighting matrix can be traced backwards into the process of formulating the pseudo inverse, which namely comes from these two equations:

$$\frac{d^TH}{dx} = 0 = x + A^Ty$$
$$\frac{d^TH}{d\lambda} = 0 = Ax - y$$

The bottom equation comes purely from our constraint ($Ax = y$, as does the second term of the first equation), so these terms cannot be influenced by the weighting B. Therefore, the weighting matrix B must come into play from the objective function alone. If we assume the objective function looks like

$$\frac{1}{2}x^TBx$$

Then the partial derivative process yields the following system:

$$\frac{d^T H}{dx} = 0 = Bx + A^T\lambda$$
$$\frac{d^T H}{d\lambda} = 0 = Ax - y$$

Due to the algebraic manipulation of these equations we performed in lecture to reach the pseudo inverse, I believe we must at some point invert the B matrix, in which case we can find equations such as

$$x = -B^{-1}A^T\lambda$$

$$Ax = -AB^{-1}A^T\lambda$$

we can plug this equation into the "constraint" equation and solve for λ in terms of y

$$(-AB^{-1}A^T\lambda) = y$$
$$\lambda = (-AB^{-1}A^T)^{-1}y$$

and finally, solving the first equation above in terms of x and y yields

$$Bx = A^T(AB^{-1}A^T)^{-1}y$$
$$x = B^{-1}A^T(AB^{-1}A^T)y$$

Obviously, this includes inverses of B instead of B itself, so I believe that means we must require an inverse of B in the objective function to then recover the given result. Therefore, this pseudo inverse given is minimizing the objective function:

$$\frac{1}{2}x^T B^{-1}x$$

where B is an invertible matrix of weightings on the states of x. This result seems a bit strange - but I'm not sure where I went wrong in my assumptions....

3 - Left Eigenvectors

Let all eigenvalues of A be distinct and let q_i to be a right eigenvector of A associated with λ_i; that is, $Aq_i = \lambda_i q_i$
Define $Q = [q_1 \quad q_2 \quad \dots \quad q_n]$ and $P = [p_1 \quad p_2 \quad \dots \quad p_n]^T$ and define

$$P = Q^{-1} = \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix}$$

Show that p_i is a left eigenvector of A associated with λ_i , that is, $p_i A = \lambda_i p_i$

We suppose the existence of an eigendecomposition of our matrix A - such a decomposition exists for any square matrix.

$$A = P\Lambda Q$$

where Λ is diagonal with scalar eigenvalues λ on the diagonal. This decomposition satisfies the "right" eigenvalue equation, as we can multiply both sides by P (AKA Q^{-1}) to recover:

$$AQ^{-1} = P\Lambda QQ^{-1}$$
$$AP = P\Lambda I$$
$$AP = P\Lambda$$

each individual equation of this system satisfies the given definition of a right eigenvector given in the problem statement.

Similarly, we can use the same method to premultiply instead by P^{-1} (AKA Q), yielding an allegorical relationship:

$$P^{-1}A = P^{-1}P\Lambda Q$$
$$P^{-1}A = I\Lambda Q$$
$$QA = \Lambda Q$$

similarly, each individual equation of this system satisfies the given definition on a left eigenvector. Note that left eigenvectors are row vectors due to the left hand multiplication of A. In the same way as the right eigenvectors, the left eigenvectors represent directions present in the rows of A that result in unrotated output when multiplied by a vector x. It is also true that the left eigenvectors of A are the right eigenvectors of the matrix A^T - This property can be shown by transposing the eigenvector equation.

4 - Null Space of the Adjoint

prove or disprove the following: (to disprove you must provide a counterexample).
Let $A : V \to U$ be a linear map, and let A^* be the adjoint of A. Then $N(A^*) = N(AA^*)$

The null space of a matrix implies the following formula holds for a given vector v. That is to say, if v is in the null space of a matrix A, then

$$Av = 0$$

similarly for the adjoint A^*:

$$A^*v = 0$$

implies that v is in the null space of A^*. No we can examine the null space of the product of these two matrices, A and A^*. For a vector v to be in this space the following must be satisfied:

$$AA^*v = 0$$

this in turn implies

$$A(A^*v) = 0$$

If we assume that v is in the null space of A^*, then we know already that $A^*v = 0$. Thus

$$A(0) = 0$$

this is now a trivial homogeneous property. Essentially the existence of the the null space of A^* implies that *any* vector falls into the null space of XA^* (where X is an arbitrary matrix of compliant dimensions), because the operation A^*v is an irreversible mapping to the zero vector.

Note that clearly, the opposite complimentary case does not hold, that is

$$N(X) \neq N(XA^*)$$

likewise, in the given example $N(A) \neq N(AA^*)$. Finding examples that prove this inequality is simple, as any value that is in the null space of A^* but not in the null space of A will not satisfy this counter-counter example. but indeed, $N(A^*) = N(AA^*)$, and no counterexamples can be found which disprove such an identity.

5 - Eigenvalues, Eigenvectors, Jordan Form

For each matrix, find the eigenvalues. For each eigenvalue, find the matrix $A - \lambda I$ and $(A - \lambda I)^k$ as necessary. Then, find the real and generalized eigenvectors, and the Jordan form. You may use MATLAB for computation if you wish.

All of these computations can be done quite quickly using MATLAB built-in functions. I assume the intention is not to rely on those alone

5.a

$$\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$$

see attached Matlab

- eigenvalues at 1, 2, 3 (distinct)
- no generalized eigenvectors

$\lambda = 1$

$$A - 1I = \begin{bmatrix} 7 & -8 & -2 \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{bmatrix}$$

eigenvector $v_1 = \begin{bmatrix} 2 & \frac{3}{2} & 1 \end{bmatrix}^T$

$\lambda = 2$

$$A - 2I = \begin{bmatrix} 6 & -8 & -2 \\ 4 & -5 & -2 \\ 3 & -4 & -1 \end{bmatrix}$$

eigenvector $v_2 = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix}^T$

$\lambda = 3$

$$A - 3I = \begin{bmatrix} 5 & -8 & -2 \\ 4 & -6 & -2 \\ 3 & -4 & -2 \end{bmatrix}$$

eigenvector $v_3 = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}^T$

Jordan form

since all the eigenvectors are distinct, we have no generalized eigenvectors and the Jordan form is just

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

5.b

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

see attached Matlab

- eigenvalues at -3, 3, 3
- no generalized eigenvectors

$\lambda = -3$

$$A + 3I = \begin{bmatrix} 4 & 0 & -4 \\ 0 & 6 & 0 \\ -2 & 0 & 2 \end{bmatrix}$$

eigenvector $v_3 = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T$

$$\lambda = 3$$

$\lambda = 3$

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

eigenvector $v_3 = \begin{bmatrix} -2 & 0 & 1 \end{bmatrix}^T$
eigenvector $v_3 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$

There are non distinct eigenvalues but all provide enough distinct eigenvectors for their multiplicity, hence, there are no generalized eigenvectors.

$$J = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

5.c

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

see attached Matlab

- eigenvalues at 2, 2, 2
- generalized eigenvector

$\lambda = 2$

$$A - 2I = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}$$

eigenvector $v_{2_1} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$
eigenvector $v_{2_2} = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}^T$

Now we have an outstanding multiplicity and not enough eigenvectors to cover it, we need to look for generalized eigenvectors in that case - so instead of solving for $Av = 0$, we can now solve for $Aw = v$ where w is a generalized eigenvector

immediately, we can see that there is no viable solutions for each of our eigenvectors, but we can also combine the baseline eigenvectors and attempt to solve the equation $Aw = v_1 + v_2$. When we consider this option, we can solve the following equation: $(A - 2I)w = v_{2_1} + v_{2_2}$

$$A - 2I = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

in this case, we can find the generalized eigenvector $w = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$ which satisfies the chaining equation.

The jordan form is thus:

$$J = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

or

$$J = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

and the largest jordan block is of size 2.