Homework 1

Exercise 2.1 To prove first we prove that the inverse of a lower (upper) matrix is lower (upper) triangular.

Take some lower triangular matrix A. Call its inverse B. Then:

$$AB = I \tag{1}$$

Writing out things by components:

$$AB = [Ab_1, \ldots, Ab_n] \tag{2}$$

So $Ab_k = e_k$. From this note that because of this we have $(Ab_k)_i = \sum_{j=1}^{i} A_{ij}(b_k)_j = 0$ For $i < k$ (Note that since A is lower triangular we only sum up to i). In particular if $i = 1 < k$ then:

$$(Ab_k)_1 = A_{11}(b_k)_1 = 0 \tag{3}$$

So for all of the collumns $k > 1$ we know that b_k

$$\begin{bmatrix} B_{11} & 0 & \ldots & 0 \\ B_{12} & \star & \star & \star \\ \ldots & \ldots & \ddots & \ldots \\ \star & \star & \star & \star \end{bmatrix} \tag{4}$$

We now prove this by induction that

$$(b_k)_1, \ldots (b_k)_{k-1} = 0 \tag{5}$$

The base case was handled above. Now assume that it is true for $1 \leq r < k$ we now prove it for $r + 1 < k$

Take

$$(Ab_k)_{r+1} = \sum_{j=1}^{n}(A_{r+1,j}(b_k)_j) \tag{6}$$

$$= \sum_{j=1}^{r+1}(A_{r+1,j}(b_k)_j) \text{ since A is lower triangular} \tag{7}$$

$$= A_{r+1,r+1}(b_k)_{r+1} \text{ induction} \tag{8}$$

$$= 0 \text{ since } (e_k)_{r+1} = 0 \tag{9}$$

However since A is invertible $A_{r+1,r+1} \neq 0$ thus b_k has zeros above the entry k.

As a result B is lower triangular. so the statement follows. Similarly if A is upper triangular then A^* is lower triangular with inverse $(A^*)^{-1}$ That is also lower triangular. Thus $(A^*)^{-1} = (A^{-1})^*$ is alos lower triangular so A^{-1} is upper triangular.

Now to prove the main theorem. If A is unitary it is diagonal.

If A is lower triangular. Then its conjugate transpose is upper triangular. But by the previous theorem its inverse is lower triangular. Since it is hermetian its conjugate transpose is its inverse so it must be both upper and lower triangular. So it must be diagonal. As a result if its conjugate transpose is diagonal then so is A

Similarly if A is upper triangular its conjugate transpose is lower triangular but its inverse is upper. As a result its inverse is both upper and lower triangular so it must be diagonal. As a result if its cnojugate transpose is diagonal then so is A.

Exercise 1.2 a) To prove this take

$$\|Ux\|_2^2 = \langle Ux, Ux \rangle \tag{10}$$
$$= x^\star U^\star U x = x^\star x = \langle x, x \rangle \tag{11}$$
$$= \|x\|_2^2 \tag{12}$$

Taking square roots of both sides yields the theorem

b)

$$\|UA\|_2 \tag{13}$$
$$= \sup_x \frac{\|UAx\|_2}{\|x\|} \tag{14}$$
$$= \sup_x \frac{\|U(Ax)\|_2}{\|x\|} \tag{15}$$
$$= \sup_x \frac{\|Ax\|_2}{\|x\|} \text{by unitary invariance on vectors} \tag{16}$$
$$= \|A\|_2 \tag{17}$$

$$\|AU\|_2 \tag{18}$$
$$= \sup_x \frac{\|AUx\|_2}{\|x\|} \tag{19}$$
$$= \sup_x \frac{\|A(Ux)\|_2}{\|U^{-1}(Ux)\|} \tag{20}$$
$$= \sup_y \frac{\|A(y)\|_2}{\|U^{-1}(y)\|} \tag{21}$$
$$= \sup_y \frac{\|A(y)\|_2}{\|(y)\|} \tag{22}$$
$$= \|A\| \tag{23}$$

2

So it is unitarily invariant.

c) Take the frobenius norm:

$$\|UA\|_F = \sqrt{\text{trace}((UA)^\star UA)} \tag{24}$$
$$= \sqrt{\text{trace}(A^\star U^\star UA)} \tag{25}$$
$$= \sqrt{\text{trace}(A^\star A)} \tag{26}$$
$$\|A\|_F \tag{27}$$

Similarly we can do:

$$\|AU\|_F = \sqrt{\text{trace}(AU(AU)^*)} \tag{28}$$
$$= \sqrt{\text{trace}(AUU^*A^*)} \tag{29}$$
$$= \sqrt{\text{trace}(AA^*)} \tag{30}$$
$$\|A\|_F \tag{31}$$

So it is unitarily invariant.

d) This is not quite the same as unitary but the same proof follows:

$$\|Wx\|_2^2 = \langle Wx, Wx \rangle \tag{32}$$
$$= x^\star W^\star W x = x^\star x = \langle x, x \rangle \tag{33}$$
$$= \|x\|_2^2 \tag{34}$$

Again taking square roots afterwards yields the theorem.

Exercise 2.3 To prove this Assume the the projector is hermitian then, Note that P projects to the range and $I - P$ projects to the null space. We show that any two vectors in this are orthogonal

$$\langle Py, (I - P)x \rangle \tag{35}$$
$$= y^* P^* (I - P)x = y^* P(I - P)x = y^* (P - P^2)x \tag{36}$$
$$= y^* (P - P)x = 0 \tag{37}$$

So the range and null space are orthogonal.

To prove the other direction. Assume that the range and null space are orthogonal. construct and orthonormal basis for the range and null space note that since the dimension of range and null space add up to the dimension of the whole space (and they are orthogonal by assumption) then this consitutes an orthonormal basis for the whole space

Basis for range $q_1, \ldots, q_k$, basis for null space $q_{k+1} \ldots q_n$.

From this note that if a vector q_j is in the range of P then:

$$Px = q_j \text{ for some x} \tag{38}$$
$$P^2 x = Pq_j \tag{39}$$
$$Px = Pq_j \tag{40}$$
$$q_j = Pq_j \tag{41}$$

3

So each vector in this basis is also an eigenvector with eigenvalue 1.

call this matrix Q then note that (if q_k) is in the range subseciton

$$Q^*PQe_k = Q^*Pq_k \tag{42}$$

$$= Q^*q_k = e_k \tag{43}$$

if q_k is in the null section of the basis:

$$Q^*PQe_k = Q^*Pq_k \tag{44}$$

$$= 0 \tag{45}$$

Thus $Q^*PQ = I_k$ where I_k is the identity up until after k after which its zeros. so since Q is unitary:

$$Q^*PQ = I_k \tag{46}$$

$$P = QI_kQ^* \tag{47}$$

$$P = Q_kQ_k^* \tag{48}$$

So P is thus hermitian since $P^* = (Q_kQ_k^*)^* = Q_kQ_k^* = P$. (Note that here Q_k is Q excluding the last $k+1, n$ collumns).

Exercise 2.4 to show this we assume By way of contradiction that there are two projection matrices. P, Q

As an aside. From this note that if a vector v is in the range of any projection P then:

$$Px = v \text{ for some x} \tag{49}$$

$$P^2x = Pv \tag{50}$$

$$Px = Pv \tag{51}$$

$$v = Pv \tag{52}$$

So each vector in this basis is also an eigenvector with eigenvalue 1.

now moving on. We show that P and Q do the exact same thing to arbitrary vectors

Take some artbirary vector x. Note that sicne the space $\mathbb{R}^n$ is a direct sum of the range and null space we can write $x = v + w$ where $v \in$ the range and $w \in$ null space specified earlier.

then note that:

$$P(x) = P(v + w) = Pv \text{ since w is in the null space of Q} \tag{53}$$

$$= v \text{ by earlier lemma} \tag{54}$$

similarly

$$Q(x) = Q(v + w) = Qv \text{ since w is in the null space of Q} \tag{55}$$

$$= v \text{ by earlier lemma} \tag{56}$$

So they do the exact same thing to arbitrary vectors. So Q and P must be the same matrix.

Exercise 2.5

Assume that $w^\star u = -1$ then we show that this is singular:

$$(I + uw^\star)u = u + uw^\star u = u + u(-1) = 0 \tag{57}$$

So $I + uw^\star$ must be singular.

Assume that it is singular then it must have a null space with some vector $v \neq 0$ in it:

$$(I + uw^\star)v = 0 \tag{58}$$
$$v = -uw^\star v \tag{59}$$
$$w^\star v = -(w^\star u)w^\star v \tag{60}$$

Now note that if $w^\star v = 0$ then $v - u(w^\star v) = 0$ a contradiction. so $w^\star v \neq 0$. and as a result

$$w^\star v = -(w^\star u)w^\star v - (w^\star u) = 1 \tag{61}$$
$$w^\star u = -1 \tag{62}$$

b)take

$$K = (I + uw^*)^{-1} - I \tag{63}$$
$$(I + uw^*)K = I - I - uw^* \tag{64}$$
$$= -uw^\star \tag{65}$$
$$\implies \tag{66}$$
$$K = (-(I + uw^\star)^{-1}u)w^\star \tag{67}$$

Note that $(-(I + uw^\star)^{-1}u) = v$ is a vector so $K = vw^\star$ so it is rank one

c) by the previous exercise (just rearranings for the inverse) we know that there is a vector v with

$$(I + uw^*)^{-1} = K + I = vw^* + I \tag{68}$$

We solve for v

$$v = (-(I + uw^\star)^{-1}u) \tag{69}$$
$$(I + uw^\star)v = (-u) \tag{70}$$
$$v + uw^*v = (-u) \tag{71}$$
$$w^*v + w^*uw^*v = (-w^*u) \tag{72}$$
$$(1 + w^*u)w^*v = (-w^*u) \tag{73}$$
$$w^*v = (-w^*u)/(1 + w^\star u) \tag{74}$$
$$\tag{75}$$

5

Since we know that $w^\star u \neq 0$

$$v + uw^\star v = -u \tag{76}$$
$$v + u(-w^*u)/(1 + w^\star u) = -u \tag{77}$$
$$v = u((w^*u)/(1 + w^\star u) - 1) \tag{78}$$
$$v = u(\frac{w^*u - 1 - w^*u}{1 + w^\star u}) \tag{79}$$
$$v = u(\frac{-1}{1 + w^\star u}) \tag{80}$$
$$\tag{81}$$

So then we know that $(i + uw^*)^{-1} = \frac{-uw^*}{1+w^*u} + I$

d) To do this take:

$$A + uw^* = A(I + A^{-1}uw^*) \tag{82}$$

Note that since A is invertible $A + uw^*$ is invertibel iff $(I + A^{-1}uw^*)$ is invertible. Thus by a previous part this is invertible iff $w^*(A^{-1}u) \neq -1$ (Since this is a rank one pertubation of the identity)

By a previous exercise we know that:

$$(I + A^{-1}uw^*)^{-1} = I - \frac{(A^{-1}uw^*)}{1 + w^*(A^{-1}u)} \tag{83}$$

$$(A + uw^*)^{-1} = (I + A^{-1}uw^*)^{-1}A^{-1} = A^{-1} - \frac{(A^{-1}uw^*)}{1 + w^*(A^{-1}u)}A^{-1} \tag{84}$$

$$\tag{85}$$

Exercise 2.6   To prove this first take some matrix:

$$P = \begin{bmatrix} 1 & \alpha & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \tag{86}$$

Note that

$$P^2 = P \tag{87}$$

So it is a projection. We calculate the largest singular value to compute its 2 norm:

$$PP^H \tag{88}$$

$$= \begin{bmatrix} 1 & \alpha & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ \overline{\alpha} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \tag{89}$$

$$= \begin{bmatrix} 1 + \|\alpha\|^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{90}$$

From here the largest singular value is clearly $\sqrt{1 + \|\alpha^2\|}$ we want this to be equal to M so:

$$M^2 = 1 + \|\alpha\|^2 \tag{91}$$
$$\alpha = \sqrt{M^2 - 1} \tag{92}$$

So choose the matrix

$$\begin{bmatrix} 1 & \sqrt{M^2 - 1} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \tag{93}$$

Exercise 2.7 To Do this let v be an eigenvalue of P then:

$$Pv = \lambda v \tag{94}$$
$$P^2 v = \lambda P v \tag{95}$$
$$Pv = \lambda P v \tag{96}$$
$$\lambda v = \lambda^2 v \tag{97}$$
$$\lambda(\lambda - 1)v = 0 \tag{98}$$

Since v has to be nonzero. any eigenvalue of this matrix will either be zero or one. We prove later that the combined total geoemtric multiplicities is the dimension of the spcae.

Construct a basis for the range of P call it $q_1, \dots, q_k$ and a basis for the null space of $q_{k+1}, \dots, q_n$. Note that putting these two basis together constructs a basis for the whole space (however this is not necessarily a unitary basis) by the rank nullity theorem.

From this note that if a vector q_j is in the range of P then:

$$Px = q_j \text{ for some x} \tag{99}$$
$$P^2 x = P q_j \tag{100}$$
$$Px = P q_j \tag{101}$$
$$q_j = P q_j \tag{102}$$

So each vector in this basis is also an eigenvector with eigenvalue 1. (So the dimension of this eigenspace is simply the dimension of the range). However the dimension of the range and null space (which nicely just so happens to be the dimension of the zero eigenspace) add up to the dimension of the whole space (by the rank nullity theorem).

Written out (E_λ is the eigenspace corresponding to eigenvalue λ)

$$\dim E_0 + \dim E_1 = \dim \text{range} + \dim \text{null} = n \tag{103}$$

Additionally the algebraic multiplicites of each eigenvalue must add up to n. but since the geometric multiplicities of each eigenvalue can never add up to n, the only way for both sums to work is if the algebraic multiplicities equal the geoemtric ones. So since the multiplicities are the same it is diagonalizable. We also show this separately below.

then take:

$$PQ = [Pq_1, \ldots, Pq_k, Pq_{k+1}, \ldots, Pq_n] \tag{104}$$
$$= [q_1, \ldots, q_k, 0, \ldots, 0] \tag{105}$$
$$= Q\Lambda \tag{106}$$

Where $\Lambda = I_k$ which is the identity up to entry k then zeros everywhere else. So this shows us an eigendecoposition $P = Q\Lambda Q^{-1}$.

So in total the eigenvectors of P correspond to vectors in the basis of the range (eigenvalue 1) and the vectors in the basis of the null space (eigenvalue 0). And the eigendecomposition is $P = Q\Lambda Q^{-1}$

Also since we didn't have any conditions on P, all projection matrices P are diagonalizable.

Exercise 2.8 Let $A = V_A\Lambda_B V_A^{-1}, B = V_B\Lambda_B V_B^{-1}$.

Call the giant matrix we were dealing with M. Then set $V = V_A \otimes V_B$ (Taking our best guess at what the eigenvectors would be) take:

$$MV = \begin{bmatrix} a_{1,1}B & a_{1,2}B & \ldots & a_{1,m}B \\ a_{2,1}B & a_{2,2}B & \ldots & a_{2,m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}B & a_{m,2}B & \ldots & a_{m,m}B \end{bmatrix} \begin{bmatrix} v_{1,1}V_B & v_{1,2}V_B & \ldots & v_{1,m}V_B \\ v_{2,1}V_B & v_{2,2}V_B & \ldots & v_{2,m}V_B \\ \vdots & \vdots & \ddots & \vdots \\ v_{m,1}V_B & v_{m,2}V_B & \ldots & v_{m,m}V_B \end{bmatrix} \tag{107}$$

$$= \begin{bmatrix} \sum_{j=1} a_{1,j}v_{j,1}BV_B & \sum_{j=1} a_{1,j}v_{j,2}BV_B & \ldots & \sum_{j=1} a_{1,j}v_{j,m}BV_B \\ \sum_{j=1} a_{2,j}v_{j,1}BV_B & \sum_{j=1} a_{2,j}v_{j,2}BV_B & \ldots & \sum_{j=1} a_{2,j}v_{j,m}BV_B \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1} a_{m,j}v_{j,1}BV_B & \sum_{j=1} a_{m,j}v_{j,2}BV_B & \ldots & \sum_{j=1} a_{m,j}v_{j,m}BV_B \end{bmatrix} \tag{108}$$

$$= \begin{bmatrix} (\sum_{j=1} a_{1,j}v_{j,1})BV_B & (\sum_{j=1} a_{1,j}v_{j,2})BV_B & \ldots & (\sum_{j=1} a_{1,j}v_{j,m})BV_B \\ (\sum_{j=1} a_{2,j}v_{j,1})BV_B & (\sum_{j=1} a_{2,j}v_{j,2})BV_B & \ldots & (\sum_{j=1} a_{2,j}v_{j,m})BV_B \\ \vdots & \vdots & \ddots & \vdots \\ (\sum_{j=1} a_{m,j}v_{j,1})BV_B & (\sum_{j=1} a_{m,j}v_{j,2})BV_B & \ldots & (\sum_{j=1} a_{m,j}v_{j,m})BV_B \end{bmatrix} \tag{109}$$

Where the lower case v in this case refers to entries of V_A and the λ_i refers to eigenvalues of A.. Note that

Now note that since each $v_{:,j}$ is an eigenvector of A with corresponding eigenvalue λ_j

$$= \begin{bmatrix} (\lambda_1 v_{1,1})BV_B & (\lambda_2 v_{1,2})BV_B & \dots & (\lambda_m v_{1,m})BV_B \\ (\lambda_1 v_{2,1})BV_B & (\lambda_2 v_{2,2})BV_B & \dots & (\lambda_m v_{2,m})BV_B \\ \vdots & \vdots & \ddots & \vdots \\ (\lambda_1 v_{m,1})BV_B & (\lambda_2 v_{m,2})BV_B & \dots & (\lambda_m v_{m,m})BV_B \end{bmatrix} \tag{110}$$

$$= \begin{bmatrix} (\lambda_1 v_{1,1})V_B\Lambda_B & (\lambda_2 v_{1,2})V_B\Lambda_B & \dots & (\lambda_m v_{1,m})V_B\Lambda_B \\ (\lambda_1 v_{2,1})V_B\Lambda_B & (\lambda_2 v_{2,2})V_B\Lambda_B & \dots & (\lambda_m v_{2,m})V_B\Lambda_B \\ \vdots & \vdots & \ddots & \vdots \\ (\lambda_1 v_{m,1})V_B\Lambda_B & (\lambda_2 v_{m,2})V_B\Lambda_B & \dots & (\lambda_m v_{m,m})V_B\Lambda_B \end{bmatrix} \tag{111}$$

$$= \begin{bmatrix} (v_{1,1})V_B\lambda_1\Lambda_B & (v_{1,2})V_B\lambda_2\Lambda_B & \dots & (v_{1,m})V_B\lambda_m\Lambda_B \\ (v_{2,1})V_B\lambda_1\Lambda_B & (v_{2,2})V_B\lambda_2\Lambda_B & \dots & (v_{2,m})V_B\lambda_m\Lambda_B \\ \vdots & \vdots & \ddots & \vdots \\ (v_{m,1})V_B\lambda_1\Lambda_B & (v_{m,2})V_B\lambda_2\Lambda_B & \dots & (v_{m,m})V_B\lambda_m\Lambda_B \end{bmatrix} \tag{112}$$

$$= \begin{bmatrix} v_{1,1}V_B & v_{1,2}V_B & \dots & v_{1,m}V_B \\ v_{2,1}V_B & v_{2,2}V_B & \dots & v_{2,m}V_B \\ \vdots & \vdots & \ddots & \vdots \\ v_{m,1}V_B & v_{m,2}V_B & \dots & v_{m,m}V_B \end{bmatrix} \begin{bmatrix} \lambda_1\Lambda_B & 0 & \dots & 0 \\ 0 & \lambda_2\Lambda_B & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m\Lambda_B \end{bmatrix} = V\Lambda \tag{113}$$

Where Λ is now diagonal. So we have expressed our matrix M as an eigendecomposition

$$MV = V\Lambda \tag{114}$$

So the eigenvector matrix is $V = V_A \otimes V_B$ with corresponding eigenvalues Λ as defined above.