Exam 2

To avoid any confusion, unless specified otherwise, vector spaces are complex vector spaces, inner-products are complex inner-products, and matrices are complex matrices. The standard inner product is $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{v}^H \boldsymbol{u} = \sum_{i=1}^n \bar{\boldsymbol{v}}_i \boldsymbol{u}_i$.

Part I: True/False

Each problem is points for a total of 50 points. (5 points each.)

You do not need to justify the answers here, this is unlike the quizzes.

1. ______ If U is unitary, then U is itself unitarily diagonalizable. This means there is a unitary V so that $U = VDV^H$ where D is diagonal.

   This is true. $U^H U = UU^H = I$, so U is normal, hence unitarily diagonalizable.

2. ______ For any diagonalizable matrix A, one can use Gram-Schmidt to find an orthogonal basis consisting of eigenvectors.

   This is false. You must first have that the eigenspaces for different eigenvalues are orthogonal.

3. ______ The collection of rank k $n \times n$ matrices is a subspace of $\mathbb{R}^{n\times n}$, for $k < n$.

   This is false, in fact SVD shows how to write any matrix as a sum of rank 1 matrices.

4. ______ If A is unitary, then $|\lambda| = 1$ for all eigenvalues λ of A.

   This is true. Let λ be an eigenvalue, with unit eigenvector $\boldsymbol{v}$. then $\langle A\boldsymbol{v}, A\boldsymbol{v}\rangle = \langle \lambda\boldsymbol{v}, \lambda\boldsymbol{v}\rangle = \bar{\lambda}\lambda\|\boldsymbol{v}\|_2^2 = |\lambda|^2 = (A\boldsymbol{v})^H(A\boldsymbol{v}) = \boldsymbol{v}^H(A^H A)\boldsymbol{v} = \boldsymbol{v}^H I \boldsymbol{v} = \|\boldsymbol{v}\|_2^2 = 1$. So $|\lambda|^2 = 1$.

5. ______ If $p(t)$ is a polynomial and $\boldsymbol{v}$ is an eigenvector of A with associated eigenvalue λ, then $p(A)\boldsymbol{v} = p(\lambda)\boldsymbol{v}$.

   This is true and trivial. $p(x) = \sum_{i=1}^k a_i x^i$, so $p(A)\boldsymbol{v} = \sum_{i=1}^k a_i A^i \boldsymbol{v} = \sum_{i=1}^k a_i \lambda^i \boldsymbol{v} = p(\lambda)\boldsymbol{v}$

6. ______ If A and B are both $n \times n$ and $\mathcal{B}$ is a basis for $\mathbb{C}^n$ consisting of eigenvectors for both A and B, then A and B commute.

   This is true. $AB = (SD_AS^{-1})(SB_BS^{-1}) = AD_AD_BS^{-1} = SD_BD_AS^{-1} = (SD_BS^{-1})(SD_AS^{-1}) = BA$.

7. ______ Any matrix A can be written as a weighted sum of rank 1 matrices..

   This is true and is essentially one of the statements of the SVD. $U\Sigma V^T = \sum_{i=1}^r \sigma_i u_i v_i^T$ where $r = \text{rank}(A)$. Each $u_i v_i^T$ is an $m \times n$ rank-1 matrix.

8. ______ For all Hermitian matrices A, there is a matrix B so that $B^H B = A$.

   This is false. A variant that is true is given in the first problem in part III. The point is that $B^H B$ is not only Hermitian, but also positive.

9. ______ There are linear maps $L : \mathbb{R}^5 \to \mathbb{R}^4$ such that $\dim(\ker(L)) = 2 = \dim(\text{rng}(L))$.

This is false, $\dim(\text{rng}(L))+\dim(\ker(L)) = \dim(\text{dom}(L))$. This is essentially the rank-nullity theorem.

10. ______ If A is invertible, then $ABA^{-1} = B$.

This is false, it would only be true if A and B commute.

Part II: Computational (60 points)

P1. (15 points) Find B so that $B^2 = A$ where

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

First diagonalize A.

Find the eigenvalues:

$\det\left(\left[\begin{smallmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{smallmatrix}\right]\right) = (1-\lambda)((2-\lambda)(1-\lambda)-1) - (-1)((-1)(1-\lambda)-0) = (1-\lambda)(1-3\lambda+\lambda^2) - (1-\lambda) = (1-\lambda)((1-3\lambda-\lambda^2-1) = (1-\lambda)(-3\lambda+\lambda^2) = (1-\lambda)(\lambda)(-3+\lambda)$. So the eigenvalues are $\lambda_1 = 3 > \lambda_2 = 1 > \lambda_3 = 0$.

This means $A = S\left[\begin{smallmatrix} 3 & & \\ & 1 & \\ & & 0 \end{smallmatrix}\right]S^{-1}$ and so $B = S\left[\begin{smallmatrix} \sqrt{3} & & \\ & 1 & \\ & & 0 \end{smallmatrix}\right]S^{-1}$ will be our matrix, where $S = \left[\begin{smallmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{smallmatrix}\right]$ where $\boldsymbol{v}_i$ is an eigenvector for λ_i.

Find eigenspaces:

$E_3 = \mathrm{NS}\left(\left[\begin{smallmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{smallmatrix}\right]\right) = \mathrm{NS}\left(\left[\begin{smallmatrix} -1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{smallmatrix}\right]\right) = \mathrm{span}\left(\left[\begin{smallmatrix} 1 \\ -2 \\ 1 \end{smallmatrix}\right]\right)$

$E_1 = \mathrm{NS}\left(\left[\begin{smallmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{smallmatrix}\right]\right) = \mathrm{NS}\left(\left[\begin{smallmatrix} -1 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right]\right) = \mathrm{span}\left(\left[\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right]\right)$

$E_0 = \mathrm{NS}(A) = \mathrm{NS}\left(\left[\begin{smallmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{smallmatrix}\right]\right) = \mathrm{span}\left(\left[\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right]\right)$

So here we could use $S = \left[\begin{smallmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ 1 & -1 & 1 \end{smallmatrix}\right]$, but in the next part we want normalized vectors, so we might as well use

$$S = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$$

this is unitary so $S^{-1} = S^T$ and finally

$$B = SDS^{-1} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ -2/\sqrt{6} & 0 & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & & \\ & 1 & \\ & & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$
$$= \frac{1}{6}\begin{bmatrix} \sqrt{3}+3 & -2\sqrt{3} & \sqrt{3}-3 \\ -2\sqrt{3} & 4\sqrt{3} & -2\sqrt{3} \\ \sqrt{3}-3 & -2\sqrt{3} & \sqrt{3}+3 \end{bmatrix}$$

Notice that B is hermitian and positive, positive hermitian matrices are like "positive real numbers", they have a positive square root, that is a positive hermitian square root. Just like $2 = \sqrt{2} \cdot \sqrt{2}$. But also $\sqrt{2}$ has another "root", namely, $2 = (-\sqrt{2}\,i)(\sqrt{2}\,i) = \bar{\lambda}\lambda$. This is the point of the next problem.

P2. (15 points) Find B so that $B^H B = A$ where A is from (1). We have the SVD from P2, so the best rank 2 approximation will be

We have already done all of the work here. Let $B = D^{1/2}S^H$ where $A = SDS^H$ just as above. So

$$B = \begin{bmatrix} \sqrt{3} & & \\ & 1 & \\ & & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

P3. (15 points) Find the best rank 2 approximation to A from (1) with respect to $\|\cdot\|_F$.

You know $\operatorname{rank}(A) = 2$ so the best rank 2 approximation of A is A, but if you just plug into the computation, you get the following:

You already have the SVD of $A = U\Sigma V^T = SDS^T$, so $U = V$ in this case and $D = \Sigma$. Now the best rank-2 approximation of A is thus (using MATLAB type notation)

$$\begin{aligned} C &= S(:,1:2)D(1:2,1:2)S^T(1:2,:) \\ &= \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ -2/\sqrt{6} & 0 \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & \\ & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} = A \end{aligned}$$

P4. (15 points) Let

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

find the characteristic polynomial and all eigenvalues, both real and complex. Explain why A is diagonalizable and compute A^{2020}. Note, I do not ask you to diagonalize A.

Find eigenvalues:

$\det\left(\left[\begin{smallmatrix} -\lambda & 0 & 1 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{smallmatrix}\right]\right) = -\lambda(-\lambda)^2 - (-1) = \lambda^3 + 1$, so the roots are $e^{i\pi} = -1$, $e^{i\frac{\pi}{3}} = \frac{\sqrt{3}}{2} + i\frac{1}{2}$, and $e^{i\frac{5\pi}{3}} = \frac{\sqrt{3}}{2} - i\frac{1}{2}$.

Compute A^{2020}:

We see $2020 = 673 \cdot 3 + 1$, so $\lambda_i^{2020} = (\lambda_i^3)^{673} \cdot \lambda_i = (-1)\lambda_i = \lambda_i$. So $S^{2020} = SD^{2020}S^{-1} = S\left[\begin{smallmatrix} \lambda_1^{2020} & & \\ & \lambda_2^{2020} & \\ & & \lambda_3^{2020} \end{smallmatrix}\right]S^{-1} = S\left[\begin{smallmatrix} -\lambda_1 & & \\ & -\lambda_2 & \\ & & -\lambda_3 \end{smallmatrix}\right]S^{-1} = -S\left[\begin{smallmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{smallmatrix}\right]S^{-1} = -A$.

Note we actually don't need to know the eigenvalues, just that $\lambda^3 = -1$.

Part III: Theory and Proofs (45 points; 15 points each)

Pick three of the following four options. If you try all four, I will grade the first three, so if this is not what you intend, then just do three, or at least make it clear which I should grade.

P1. Let S be a fixed invertible $n \times n$ matrix. Let U be the set of $n \times n$ matrices that are diagonalized by S, that is $A = SD_AS^{-1}$ for some diagonal matrix A. Either prove that that U is a subspace of $\mathbb{C}^{n\times n}$ or show that U is not a subspace of $\mathbb{C}^{n\times n}$.

This is a subspace, let $A, B \in U$, so $A = SD_AS^{-1}$ and $B = SD_BS^{-1}$, so $\alpha A + B = \alpha(SD_AS^{-1}) + SD_BS^{-1} = S(\alpha D_A + D_B)S^{-1}$, so $\alpha A + B \in U$. Thus U is a subspace.

P2. Let A be a real $m \times n$ matrix and let $A^\dagger = V\Sigma^\dagger U^T$, where $A = U\Sigma V^T$ where U is $m \times m$, V is $n \times n$, both unitary, Σ is $m \times n$ and $\Sigma^\dagger$ is $n \times m$ have the form

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} \quad \text{and} \quad \Sigma^\dagger = \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

Show: $\hat{\boldsymbol{x}} = A^\dagger \boldsymbol{b}$ is a least-squares solution to $A\boldsymbol{x} = \boldsymbol{b}$.

Previously we used $\hat{\boldsymbol{x}} = (A^TA)^{-1}A^T\boldsymbol{b}$ for our least-squares solution, but we had the restriction that the columns of the "data" matrix A were independent, this guarantees that $\mathrm{NS}(A) = \mathrm{NS}(A^TA) = \{\boldsymbol{0}\}$. It is not hard to see that $A^\dagger = (A^TA)^{-1}A^T$ if A has linear independent columns.

Review the comments about Topic 5 DQ 2 in the Class Notes. Particularly point (2.) concerning what it means to be a least-squares solution to $A\boldsymbol{x} = \boldsymbol{b}$.

This was actually a homework problem, we need to show that

$$A^TA\hat{\boldsymbol{x}} = A^T\boldsymbol{b}$$

so that is

$$A^TAA^\dagger = A^T$$

Here we just compute:

$$(V\Sigma^TU^T)(U\Sigma V^T)(V\Sigma^\dagger U^T) = V\Sigma^T\Sigma\Sigma^\dagger U^T = V\Sigma^TU^T = A^T$$

The only point here is $\Sigma^T\Sigma\Sigma^\dagger = \Sigma^T$. Note sizes, Σ is $m \times n$, $\Sigma^\dagger$ is $n \times m$, and $\Sigma\Sigma^\dagger = \begin{bmatrix} I_r & \\ & 0_{m-r} \end{bmatrix}$ so $\Sigma^T(\Sigma\Sigma^\dagger) = \Sigma^T$.

Read more on the Moore-Penrose inverse here.

P3. Prove that any complex inner-product $\langle \cdot, \cdot \rangle_V$ on a complex vector space V, there is a basis $\mathcal{U} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ so that

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_V = [\boldsymbol{y}]_{\mathcal{U}}^H [\boldsymbol{x}]_{\mathcal{U}}$$

In other words for any finite dimensional inner-product space, there is a choice of basis, so that with respect to that basis, the inner-product is represented by the standard inner-product.

Here, in case you need it, is the definition of an inner-product. All the notation here is as I always use it in my notes.

Gram-Schmidt will produce an orthonormal basis for V, say $\mathcal{U} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ and then if $[\boldsymbol{x}]_{\mathcal{U}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ and $[\boldsymbol{y}]_{\mathcal{U}} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$, then

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{y} \rangle_V &= \left\langle \sum_i \alpha_i \boldsymbol{u}_i, \sum_j \beta_j \boldsymbol{u}_j \right\rangle \\
&= \sum_i \alpha_i \sum_j \bar{\beta}_j \langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle \\
&= \sum_i \alpha_i \sum_j \bar{\beta}_j \delta_{i,j} && (\delta_{i,j} = 1 \text{ if } i = j; \text{ 0 otherwise}) \\
&= \sum_i \alpha_i \bar{\beta}_i \\
&= \begin{bmatrix} \bar{\beta}_1 & \cdots & \bar{\beta}_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \\
&= [\boldsymbol{y}]_{\mathcal{U}}^H [\boldsymbol{x}]_{\mathcal{U}}
\end{aligned}$$

so

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_V = [\boldsymbol{y}]_{\mathcal{U}}^H [\boldsymbol{x}]_{\mathcal{U}}$$

as required.

P4. Use the SVD to show that any square matrix A can be written as $A = UP$ where U is unitary and P is Hermitian.

Let $A = V\Sigma W^H$ as in SVD and let $U = VW^H$, this is unitary since both V and W are unitary. So

$$A = (VW^H(W\Sigma W^H)) = UP$$

where $P = W\Sigma W^H$. This P is clearly Hermitian.

Note: If we think of matrices like generalizations of complex numbers, then Hermitian matrices are like the reals, $P^H = P$ is like $\bar{z} = z$ for $z \in \mathbb{C}$. A unitary is "like" a rotation, so here we represent A as a rotation followed by a "real." this is like writing $z = e^{i\theta} r$, the polar form of a complex number.