Final Exam

06/17/2020

1.

$r = m = n,\quad \mathbf{A}_1$ is any invertible square matrix

$r = m < n,\quad \mathbf{A}_2$ has extra columns like $\mathbf{A}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

$r = n < m,\quad \mathbf{A}_3$ has extra rows like $\mathbf{A}_2^{\,T}$

$r < m, r < n,\ \mathbf{A}_4 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{or} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \text{or} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$

2. ($\mathbf{F}^{-1}$ 설명없으면 −2점)

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \rightarrow \lambda = 1, i, i^2, i^3 \rightarrow \mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}$$

$$\rightarrow \mathbf{Q} = \frac{\mathbf{F}}{2} \rightarrow \text{orthonormal matrix} \rightarrow \mathbf{Q}\overline{\mathbf{Q}}^T = \mathbf{I} \rightarrow \mathbf{F}\overline{\mathbf{F}}^T = 4\mathbf{I}$$

$$\mathbf{F}\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \rightarrow \mathbf{x} = \mathbf{F}^{-1} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{4}\overline{\mathbf{F}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{bmatrix}$$

3. (eigenvalue정의 활용시 −1점)

$$\left.\begin{aligned} \text{tr}(\mathbf{A}) &= \sum \lambda_i \\ \det(\mathbf{A}) &= \prod \lambda_i \end{aligned}\right\} \rightarrow \mathbf{A} = \begin{bmatrix} 0 & 1 \\ -28 & 11 \end{bmatrix}$$

4. (incorrect eigenvector: 2 pts)

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \xrightarrow{\lambda=1,3} \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}\mathbf{x}_1 = \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix}\mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \rightarrow \mathbf{A}^3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 27 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \mathbf{A}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

5. (1 pt each)

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} & & \\ & S & \\ & & \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4(x_1 - x_2 + 2x_3)^2 \rightarrow S = \begin{bmatrix} 4 & -4 & 8 \\ -4 & 4 & -8 \\ 8 & -8 & 16 \end{bmatrix} \rightarrow \begin{cases} pivot = 4 \\ rank = 1 \\ eigenvalues = 0,0,24 \\ \text{determinant} = 0 \end{cases}$$

6.

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \xrightarrow{\lambda=25,0} \begin{bmatrix} -20 & 10 \\ 10 & -5 \end{bmatrix}\mathbf{v}_1 = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \mathbf{v}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$u_1 = \frac{\mathbf{A}v_1}{\sigma_1} = \frac{1}{5}\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \rightarrow \mathbf{A} = u_1\sigma_1 v_1^T = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}(5)\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{A} = \mathbf{Q}\Lambda\mathbf{Q}^T = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 25 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

(alternatively, but not preferrable)

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix} \xrightarrow{\lambda=25,0} \begin{bmatrix} -5 & 10 \\ 10 & -20 \end{bmatrix}\mathbf{u}_1 = \begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix}\mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

7. The singular values of $\mathbf{A} - \mathbf{A}_k$ are $\sigma_{k+1} \geq \sigma_{k+2} \geq \cdots \geq \sigma_r$ (the smallest $(r-k)$ singular values of $\mathbf{A}$)

(rank 언급 없을 경우 -1점)

8. (5 pts each)

(1) (m≧n) There are n nonzero in $\Sigma = \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{bmatrix}}_{m\times n}$.

(2) $$\Sigma^T\Sigma = \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{bmatrix}}_{n\times m} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{bmatrix}}_{m\times n} = \underbrace{\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix}}_{n\times n} \to \left(\Sigma^T\Sigma\right)^{-1} = \underbrace{\begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_n^2 \end{bmatrix}}_{\substack{n\times n \\ \text{invertible (nonzero diagonals)}}}$$

(3) $$\Sigma^+ = \left(\Sigma^T\Sigma\right)^{-1}\Sigma^T = \underbrace{\begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & 1/\sigma_n^2 \end{bmatrix}}_{n\times n} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{bmatrix}}_{n\times m} = \underbrace{\begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & 0 \\ & & 1/\sigma_n & \end{bmatrix}}_{n\times m}$$

(4) $$\mathbf{A}^+ = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T = \left[\left(\mathbf{U}\Sigma\mathbf{V}^T\right)^T\mathbf{U}\Sigma\mathbf{V}^T\right]^{-1}\left(\mathbf{U}\Sigma\mathbf{V}^T\right)^T = \left(\mathbf{V}\Sigma^T\underbrace{\mathbf{U}^T\mathbf{U}}_{\mathbf{I}}\Sigma\mathbf{V}^T\right)^{-1}\mathbf{V}\Sigma^T\mathbf{U}^T$$
$$= \underbrace{\left(\mathbf{V}^T\right)^{-1}}_{\mathbf{V}^T\mathbf{V}=\mathbf{I}}\left(\Sigma^T\Sigma\right)^{-1}\underbrace{\mathbf{V}^{-1}\mathbf{V}}_{\mathbf{I}}\Sigma^T\mathbf{U}^T = \mathbf{V}\left(\Sigma^T\Sigma\right)^{-1}\Sigma^T\mathbf{U}^T = \mathbf{V}\Sigma^+\mathbf{U}^T$$

9. (4/3/3 pts)

$$\mathbf{C} = \mathbf{I} + \mathbf{P} + \mathbf{P}^2 + \mathbf{P}^3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \mathbf{u}\mathbf{v}^T$$

$$\to rank(\mathbf{C}) = 1 \to \lambda = 0,0,0,4$$

$$\mathbf{Fc} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_0(\mathbf{C}) \\ \lambda_1(\mathbf{C}) \\ \lambda_2(\mathbf{C}) \\ \lambda_3(\mathbf{C}) \end{bmatrix} \to \begin{cases} 1 + i + i^2 + i^3 = 0 \to z = i \\ 1 + i^2 + i^4 + i^6 = 0 \to z = i^2 = -1 \\ 1 + i^3 + i^6 + i^9 = 0 \to z = i^3 = -i \end{cases}$$

10. loss functions

$$\begin{cases} square(L_2) : \sum \| \ \|_2^2 \rightarrow regression \\ L_1 \\ hinge : SVM \rightarrow classification(0/1) \\ cross-entropy : (soft\max) - \sum \ln P \rightarrow classification \end{cases}$$

11. (1 pt each)

output = (N+2P-F)/stride +1: N=32, P=2, F=5, stride=1 → output=32

max-pooing: 2x2 filters and stride 2 (pad 0): N → N/2

Layer	Activation volume	No. of Parameters
Input	32x32x1	0
Conv5-10	32x32x10	10x1x5x5 + 10
Pool-2	16x16x10	0
Conv5-10	16x16x10	10x10x5x5 + 10
Pool2	8x8x10	0
FC-N	10	8x8x10x10 + 10

12. (2/6/2 pts)

$$\frac{\partial L}{\partial z_1} = \begin{cases} \delta_1 \ \text{ if } z_1 = \max\{z_1, z_2, 0\} \\ 0 \ \text{ otherwise} \end{cases}$$

$$\frac{\partial L}{\partial z_2} = \begin{cases} \delta_1 + \delta_2 \ \text{ if } z_2 = \max\{z_1, z_2, 0\} \text{and } z_2 = \max\{z_2, z_3, 0\} \\ \delta_1 \ \text{ else if } z_2 = \max\{z_1, z_2, 0\} \\ \delta_2 \ \text{ else if } z_2 = \max\{z_2, z_3, 0\} \\ 0 \ \text{ otherwise} \end{cases}$$

$$\frac{\partial L}{\partial z_3} = \begin{cases} \delta_2 \ \text{ if } z_3 = \max\{z_2, z_3, 0\} \\ 0 \ \text{ otherwise} \end{cases}$$