Final Exam

06/17/2020

1. Four possibilities for the rank r and size m, n match four possibilities for $\mathbf{Ax} = \mathbf{b}$. Find four matrices $\mathbf{A}_1$ to $\mathbf{A}_4$ that show those possibilities: (10 pts)

$r = m = n,$	$\mathbf{A}_1\mathbf{x} = \mathbf{b}$ has 1 solution for every $\mathbf{b}$
$r = m < n,$	$\mathbf{A}_2\mathbf{x} = \mathbf{b}$ has 1 or ∞ solutions
$r = n < m,$	$\mathbf{A}_3\mathbf{x} = \mathbf{b}$ has 0 or 1 solution
$r < m, r < n,$	$\mathbf{A}_4\mathbf{x} = \mathbf{b}$ has 0 or ∞ solutions

2. Find eigenvectors and eigenvalues of the permutation matrix $\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

   The eigenvectors are columns of the 4 by 4 Fourier matrix $\mathbf{F}$.

   Show that $\mathbf{Q} = \mathbf{F}/2$ has orthogonal columns.

   What vector $\mathbf{x}$ has $\mathbf{Fx} = (1, 0, 1, 0)$? (10 pts)

3. Choose the last rows of $\mathbf{A}$ to give eigenvalues 4 and 7: $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ * & * \end{bmatrix}$ (5 pts)

4. Factor the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ into $\mathbf{A} = \mathbf{X\Lambda X}^{-1}$: $\mathbf{A}^3 = (\quad)(\quad)(\quad)$ and $\mathbf{A}^{-1} = (\quad)(\quad)(\quad)$ (5 pts)

5. Find the 3 by 3 matrix $\mathbf{S}$ and its pivots, rank, eigenvalues, and determinant: (5 pts)

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} & & \\ & S & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4(x_1 - x_2 + 2x_3)^2$$

6. Find the SVD of the rank 1 matrix $\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$. Factor $\mathbf{A}^T\mathbf{A}$ into $\mathbf{Q\Lambda Q}^T$. (5 pts)

7. What are the singular values (in descent order) of $\mathbf{A} - \mathbf{A}_k$? Omit any zeros. (5 pts)

8. Suppose A has independent columns (rank $r = n$; null space = zero vector) (20 pts)

(1) Describe the m by n matrix Σ in $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$. How many nonzeros in Σ?

(2) Show that $\Sigma^T\Sigma$ is invertible by finding its inverse.

(3) Write down the n by m matrix $\left(\Sigma^T\Sigma\right)^{-1}\Sigma^T$ and identify it it as Σ^+.

(4) Substitute $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ into $\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T$ and identify that matrix as $\mathbf{A}^+$.

9. What are the eigenvalue of the 4 by 4 circulant $\mathbf{C} = \mathbf{I} + \mathbf{P} + \mathbf{P}^2 + \mathbf{P}^3$?

Connect those eigenvalues to the discrete transform $\mathbf{Fc}$ for $\mathbf{c} = (1, 1, 1, 1)$.

For which three real or complex numbers z is $1 + z + z^2 + z^3 = 0$? (10 pts)

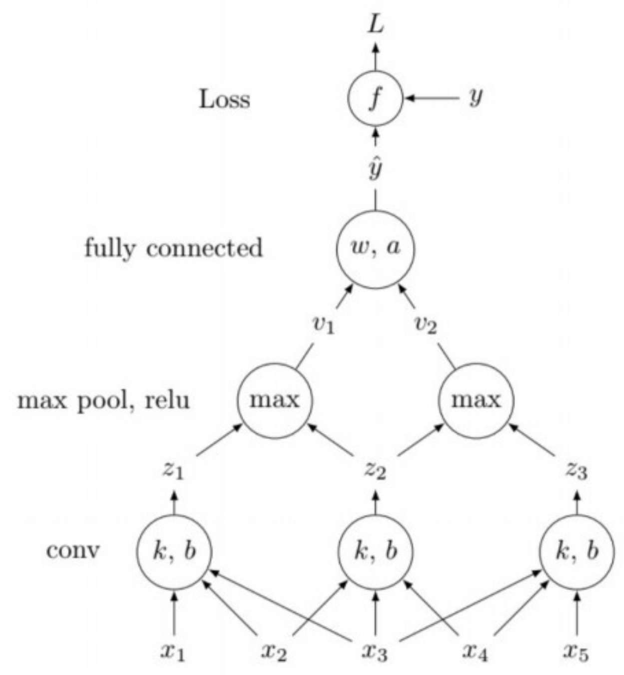
10. Classify the loss functions and describe those characteristics. (5 pts)

11. Consider the convolutional network defined by the layers in the left column below. Fill in the size of the activation volumes at each layer, and the number of parameters at each layer. You can write your answer as a multiplication (e.g. 128x128x3). (10 pts)

- CONV5-N denotes a convolutional layer with N filters, each of size 5x5xD, where D is the depth of the activation volume at the previous layer. Padding is 2 and stride is 1.
- POOL2 denotes a 2x2 max-pooling layer with stride 2 (pad 0)
- FC-N denotes a fully-connected layer with N output neurons.

Layer	Activation Volume Dimensions (memory)	Number of parameters
INPUT	32x32x1	0
CONV5-10		
POOL2		
CONV5-10		
POOL2		
FC-10		

12. Consider the following 1-dimensional ConvNet, where all variables are scalars:



$$L = \frac{1}{2}(y - \hat{y})^2$$

$$\hat{y} = \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + a$$

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \max\{z_1, z_2, 0\} \\ \max\{z_2, z_3, 0\} \end{bmatrix}$$

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & k_3 & 0 & 0 \\ 0 & k_1 & k_2 & k_3 & 0 \\ 0 & 0 & k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} b \\ b \\ b \end{bmatrix}$$

Given the gradients of the loss L with respect to the second layer activations v, derive the gradient of the loss with respect to the first layer activations z. More precisely, given $\frac{\partial L}{\partial v_i} = \delta_i$ determine $\frac{\partial L}{\partial z_i}$ (10 pts)