UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering

ECE 417 Multimedia Signal Processing
Spring 2023

EXAM 1

Tuesday, September 26, 2023

- This will be a CLOSED BOOK exam.
- You are permitted one sheet of handwritten notes, 8.5x11.
- Calculators and computers are not permitted.
- Don't simplify explicit numerical expressions.
- If you're taking the exam online, you will need to have your webcam turned on. Your exam will appear on Gradescope at exactly 9:30am; you will need to photograph and upload your answers by exactly 11:00am.
- There are a total of 100 points in the exam. Each problem specifies its point total. Plan your work accordingly.
- You must SHOW YOUR WORK to get full credit.

Name: ______________________________

NetID: ______________________________

Linear Algebra: If $\mathbf{A}$ is tall and thin, with full column rank, then

$$\mathbf{A}^{\dagger}\mathbf{b} = \text{argmin}_{\mathbf{v}}\|\mathbf{b} - \mathbf{A}\mathbf{v}\|^2 = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$$

If $\mathbf{A}$ is short and fat, with full row rank, then

$$\mathbf{A}^{\dagger} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$$

Orthogonal projection of $\mathbf{x}$ onto the columns of $\mathbf{A}$ is $\mathbf{x}_{\perp} = \mathbf{A}\mathbf{A}^{\dagger}\mathbf{x}$. Orthogonal projection onto the rows of $\mathbf{A}$ is $\mathbf{x}_{\perp} = \mathbf{A}^{\dagger}\mathbf{A}\mathbf{x}$. **Image Interpolation**

$$\begin{aligned}
y[n_1, n_2] &= \begin{cases} x\left[\frac{n_1}{U}, \frac{n_2}{U}\right] & \frac{n_1}{U}, \frac{n_2}{U} \text{ both integers} \\ 0 & \text{otherwise} \end{cases} \\
z[n_1, n_2] &= h[n_1, n_2] * y[n_1, n_2] \\
h_{\text{rect}}[n_1, n_2] &= \begin{cases} 1 & 0 \leq n_1, n_2 < U \\ 0 & \text{otherwise} \end{cases} \\
h_{\text{tri}}[n] &= \begin{cases} \left(1 - \frac{|n_1|}{U}\right)\left(1 - \frac{|n_2|}{U}\right) & -U \leq n_1, n_2 \leq U \\ 0 & \text{otherwise} \end{cases} \\
h_{\text{sinc}}[n_1, n_2] &= \frac{sin(\pi n_1/U)}{\pi n_1/U}\frac{sin(\pi n_2/U)}{\pi n_2/U}
\end{aligned}$$

Barycentric Coordinates

$$\begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3 = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

DTFT, DFT, STFT

$$\begin{aligned}
X(\omega) &= \sum_n x[n]e^{-j\omega n} \\
x[n] &= \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{j\omega n}d\omega \\
X[k] &= \sum_{n=0}^{N-1} x[n]e^{-j\frac{2\pi kn}{N}} \\
x[n] &= \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j\frac{2\pi kn}{N}} \\
X_m(\omega) &= \sum_n w[n-m]x[n]e^{-j\omega(n-m)}, \quad \omega_k = \frac{2\pi k}{N} \\
x[n] &= \frac{\sum_m \frac{1}{N}\sum_{k=0}^{N-1} X_m[k]e^{j\omega_k(n-m)}}{\sum_m w[n-m]}
\end{aligned}$$

Griffin-Lim

$$\begin{aligned}
X_t[k] &\leftarrow \text{STFT}\left\{\text{ISTFT}\left\{X_t[k]\right\}\right\} \\
X_t[k] &\leftarrow M_t[k]e^{j\angle X_t[k]}
\end{aligned}$$

1. (25 points) The extreme learning machine (XLM) is a type of two-layer neural network trained without gradient descent. The first-layer weights are sampled from a unit-normal probability density, and never trained; the second-layer weights are chosen in one step to minimize the mean-squared error with which XLM approximates its training targets. Let $\mathbf{x}_i$ be the i^{th} input vector, let $\mathbf{a}_i$ be the corresponding vector of hidden node activations, and let $\mathbf{y}_i$ be the corresponding output target vector. Find the second-layer weight matrix, $\mathbf{W}_2$, that minimizes the following mean-squared error term:

$$\mathcal{L} = \sum_{i=1}^{n} \|\mathbf{y}_i - \mathbf{W}_2^T \mathbf{a}_i\|_2^2$$

2. (25 points) You are trying to synthesize a video frame $T(\mathbf{x})$, $\mathbf{x} = [x_1, x_2]^T$ using the method of barycentric coordinates. In this video frame, the corners of one triangle are the points $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ given by

$$\mathbf{x}_1 = \begin{bmatrix} 45 \\ 15 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 30 \\ 41 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 49 \\ 82 \end{bmatrix}$$

You are trying to find the color of the pixel $\mathbf{x} = [40, 36]^T$.

(a) As a first step, notice that $\mathbf{x} = \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3$. Write an equation for β_1, β_2, and β_3 in the form of a matrix inverse times a vector. Do not try to evaluate the matrix inverse: just specify the contents of the matrix to be inverted.

(b) Suppose that, using the method in part (a) plus some other steps, you determine that the point $T(\mathbf{x})$ in the target image should have the same color as the point $S(\mathbf{y})$ in a source image $S(\cdot)$, where $\mathbf{y} = [13.6, 18.2]^T$. Unfortunately, the source image $S(\mathbf{y})$ is only specified for integer values of y_1 and y_2. Express $S([13.6, 18.2]^T)$ as a weighted sum of pixels $S(\mathbf{y})$ for which y_1 and y_2 are integers, and specify the values of the weights, using the method of bilinear interpolation.

3. (25 points) Suppose that $x[n] = \cos(\omega_0 n)$. Find its STFT, $X_m[k]$. Your answer should be expressed in terms of ω_0, m, k, the DFT length N, and the DTFT of the window, $W\left(\frac{2\pi k}{N}\right)$.

4. (25 points) Consider applying the Griffin-Lim algorithm to reconstruct $x[n]$ from $X_{tL/2}[k]$, an STFT that assumes an L-sample Hann window and a hop length of $L/2$ samples. In the first iteration, $x[n]$ is synthesized as

$$x[n] = \text{ISTFT}\left(X_{tL/2}[k]\right)$$

Notice that, by this definition, samples $x[L/2]$ through $x[L-1]$ depend on both $X_0[k]$ and $X_{L/2}[k]$. Express $x[n]$, for $\frac{L}{2} \leq n \leq L-1$, as a function of $X_0[k]$ and $X_{L/2}[k]$.

THIS PAGE IS SCRATCH PAPER.