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

ECE 417 Multimedia Signal Processing
Fall 2023

PRACTICE EXAM 1

Exam will be Tuesday, September 26, 2023

- This will be a CLOSED BOOK exam.
- You will be 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 will be 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**

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

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. (16 points) A 200×200 sunset image is bright on the bottom, and dark on top, thus the pixel in the i^{th} row and j^{th} column has intensity $A[i,j] = 200 - i$. Suppose that this formula extends past the edge of the image, and also extends to non-integer pixel locations, so you can assume that it holds for every real-valued coordinate value (i,j).

   Suppose your goal is to recognize sunset images that have been taken by people laying upside-down on the beach. You decide to train the classifier by rotating the image $A[i,j]$ to every possible angle, thus creating the training images

$$B_k[i,j] = A[i\cos\theta_k - j\sin\theta_k, i\sin\theta_k + j\cos\theta_k], \quad \theta_k = \frac{2\pi k}{100}, \quad 0 \le k \le 99$$

   Your next step is to reshape each 200×200 image $B_k[i,j]$ into a vector of raw pixel intensities, $\mathbf{x}_k$, then to compute the dataset mean, $\mathbf{m} = \frac{1}{100}\sum_{k=0}^{99}\mathbf{x}_k$.

   (a) What is the length of the vector $\mathbf{m}$?

   Solution:

$$\text{len}(\mathbf{m}) = \text{len}(\mathbf{x}_k) = 200 \times 200 = 40,000$$

   (b) What is the numerical value of $\mathbf{m}$? Provide enough information to specify the value of every element of the vector.

   Solution:

$$\begin{aligned} M[i,j] &= \frac{1}{100}\sum_{k=0}^{99} B_k[i,j] \\ &= \frac{1}{100}\sum_{k=0}^{99} A\left[i\cos\frac{2\pi k}{100} - j\sin\frac{2\pi k}{100}, i\sin\frac{2\pi k}{100} + j\cos\frac{2\pi k}{100}\right] \\ &= \frac{1}{100}\sum_{k=0}^{99}\left(200 - i\cos\frac{2\pi k}{100} + j\sin\frac{2\pi k}{100}\right) \\ &= 200 \end{aligned}$$

   Therefore

$$\mathbf{m} = [200, 200, \ldots, 200]^T$$

2. (16 points) Suppose you have a 1000-sample audio waveform, $x[n]$, such that $x[n] \neq 0$ for $0 \leq n \leq 999$. You want to chop this waveform into 200-sample frames, with 10% overlap between frames. How many nonzero samples are there in the last frame?

> **Solution:** 10% overlap = 20-sample overlap, so the frames start at samples 0, 180, 360, 540, 720, and 900. The last frame has 100 nonzero samples.

3. (21 points) You are given a 640x480 B/W input image, $x[n_2, n_1]$ for integer pixel values $0 \leq n_1 \leq 639$, $0 \leq n_2 \leq 479$. You wish to interpolate the given pixel values in order to find the value of the image at location $(n_1, n_2) = (500.3, 300.8)$. Specify the formula used to calculate $x[300.8, 500.3]$ using each of the following algorithms. Be certain that your formula clearly states which pixels from the input image are used.

   (a) Piece-wise constant interpolation.

   > **Solution:**
   >
   > $$x[300.8, 500.3] = x[300, 500]$$

   (b) Bilinear interpolation.

   > **Solution:**
   >
   > $$\begin{aligned} x[300.8, 500.3] &= (0.7)(0.2)x[300, 500] + (0.7)(0.8)x[301, 500] + \\ &\quad (0.3)(0.2)x[300, 501] + (0.3)(0.8)x[301, 501] \end{aligned}$$

   (c) Sinc interpolation.

   > **Solution:**
   >
   > $$x[300.8, 500.3] = \sum_{n_1=0}^{639} \sum_{n_2=0}^{479} x[n_2, n_1] \mathrm{sinc}\left(\pi(500.3 - n_1)\right) \mathrm{sinc}\left(\pi(300.8 - n_2)\right)$$

4. (10 points) Suppose a particular image has the following pixel values:

$$a[0,0] = 1, \quad a[1,0] = 0, \quad a[0,1] = 0, \quad a[1,1] = 0$$

Use bilinear interpolation to estimate the value of the pixel $a\left(\frac{1}{3}, \frac{1}{3}\right)$.

Solution: $a\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{4}{9}$

5. (20 points) Image warping has moved input pixel $i(4.6, 8.2)$ to output pixel $i'(15, 7)$. Input pixel $i(4.6, 8.2)$ is unknown, but you know that $i(4, 8) = a$, $i(4, 9) = b$, $i(5, 8) = c$, and $i(5, 9) = d$. Use bilinear interpolation to estimate $i(4.6, 8.2)$ in terms of a, b, c, and d.

Solution:

$$i(4.6, 8.2) = (0.4)(0.8)a + (0.4)(0.2)b + (0.6)(0.8)c + (0.6)(0.2)d$$

6. (17 points) Your goal is to find a positive real number, a, so that $ax[n]$ is as similar as possible to $y[n]$ in the sense that it minimizes the following error:

$$\epsilon = \int_{-\pi}^{\pi} \left(|Y(\omega)| - a|X(\omega)|\right)^2 d\omega$$

Find the value of a that mininimizes ϵ, in terms of $|X(\omega)|$ and $|Y(\omega)|$.

Solution:

$$\frac{\partial \epsilon}{\partial a} = 2\int_{-\pi}^{\pi} \left(a|X(\omega)| - |Y(\omega)|\right)|X(\omega)|d\omega$$

which equals 0 at:

$$a = \frac{\int_{-\pi}^{\pi} |X(\omega)||Y(\omega)|d\omega}{\int_{-\pi}^{\pi} |X(\omega)|^2 d\omega}$$

7. (5 points) Consider the problem of upsampling, by a factor of 2, the infinite-sized image

$$x[n_1, n_2] = \delta[n_1 - 5] = \begin{cases} 1 & n_1 = 5 \\ 0 & \text{otherwise} \end{cases}$$

Suppose that the image is upsampled, then filtered, as

$$y[n_1, n_2] = \begin{cases} x[n_1/2, n_2/2] & n_1/2 \text{ and } n_2/2 \text{ both integers} \\ 0 & \text{otherwise} \end{cases} \qquad z[n_1, n_2] = y[n_1, n_2] ** h[n_1, n_2]$$

Let $h[n_1, n_2]$ be the ideal anti-aliasing filter with frequency response

$$H(\omega_1, \omega_2) = \begin{cases} 1 & |\omega_1| < \frac{\pi}{2}, \quad |\omega_2| < \frac{\pi}{2} \\ 0 & \text{otherwise} \end{cases}$$

Find $z[n_1, n_2]$.

Solution:

$$h[n_1, n_2] = h_1[n_1]h_2[n_2] = \left(\frac{1}{2}\right) \text{sinc}\left(\frac{\pi n_1}{2}\right) \left(\frac{1}{2}\right) \text{sinc}\left(\frac{\pi n_2}{2}\right)$$

$$y[n_1, n_2] = \begin{cases} 1 & n_1 = 10 \text{ and } n_2 \text{ a multiple of } 2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \left(\sum_{p=-\infty}^{\infty} \delta[n_2 - 2p]\right) & n_1 = 10 \\ 0 & \text{otherwise} \end{cases}$$

Convolving along each row gives $h_2[n_2] * y[n_1, n_2]$, which is zero, except on the $n_1 = 10$ row. On that row, $y[n_1, n_2]$ is equal to one on the even-numbered samples, and equal to zero on the odd-numbered samples. The correct answer is the obvious one: the low-pass filter computes a perfect average between 0 and 1, so each pixel winds up with a value of 1/2. If you want to do a more careful analysis, you could notice that this row is an impulse train with a period of $P = 2$, and therefore it has a DTFT which has impulses of area $2\pi/P = \pi$ at $\omega = 0$ and $\omega = \pi$.

The LPF keeps only the $\omega = 0$ impulse, thus:

$$
\begin{aligned}
h_2[n_2] * y[n_1, n_2] &= \begin{cases} \left(\sum_{p=-\infty}^{\infty} \delta[n_2 - 2p]\right) * \left(\frac{1}{2}\text{sinc}\left(\frac{\pi n_2}{2}\right)\right) & n_1 = 10 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \mathcal{F}^{-1}\left\{\left(\frac{2\pi}{2}\sum_{k=0}^{1} \delta\left(\omega - \frac{2\pi k}{2}\right)\right)\left(\begin{cases} 1 & |\omega_2| < \frac{\pi}{2} \\ 0 & \text{otherwise} \end{cases}\right)\right\} & n_1 = 10 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \mathcal{F}^{-1}\left\{\pi\delta(\omega)\right\} & n_1 = 10 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{2} & n_1 = 10 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

Convolving along each column, then, gives

$$
z[n_1, n_2] = h_1[n_1] * h_2[n_2] * y[n_1, n_2] = \left(\frac{1}{4}\right) \text{sinc}\left(\frac{\pi(n_1 - 10)}{2}\right)
$$

8. (5 points) Suppose you have a 200×200-pixel image that is just one white dot at pixel $(45, 25)$, and all the other pixels are black:

$$x[n_1, n_2] = \begin{cases} 255 & n_1 = 45,\ n_2 = 25 \\ 0 & \text{otherwise},\ 0 \le n_1 < 199,\ 0 \le n_2 < 199 \end{cases}$$

This image is upsampled to size 400×400, then filtered, as

$$y[n_1, n_2] = \begin{cases} x[n_1/2, n_2/2] & n_1/2 \text{ and } n_2/2 \text{ both integers} \\ 0 & \text{otherwise} \end{cases} \qquad z[n_1, n_2] = y[n_1, n_2] ** h[n_1, n_2]$$

where $h[n_1, n_2]$ is the ideal anti-aliasing filter whose frequency response is

$$H(\omega_1, \omega_2) = \begin{cases} 1 & |\omega_1| < \frac{\pi}{2},\ |\omega_2| < \frac{\pi}{2} \\ 0 & \text{otherwise} \end{cases}$$

Find $z[n_1, n_2]$.

Solution:

$$y[n_1, n_2] = \begin{cases} 255 & n_1 = 90, n_2 = 50 \\ 0 & \text{otherwise} \end{cases}$$

$$h[n_1, n_2] = \frac{1}{2}\text{sinc}\left(\frac{\pi n_1}{2}\right)\frac{1}{2}\text{sinc}\left(\frac{\pi n_2}{2}\right) = h_1[n_1]h_2[n_2]$$

Row convolution gives $v[n_1, n_2] = h_2[n_2] * y[n_1, n_2]$, which is

$$v[n_1, n_2] = \begin{cases} \frac{255}{2}\text{sinc}\left(\frac{\pi(n_2-50)}{2}\right) & n_1 = 90 \\ 0 & \text{otherwise} \end{cases}$$

Column convolution then gives

$$z[n_1, n_2] = \frac{255}{4}\text{sinc}\left(\frac{\pi(n_1 - 90)}{2}\right)\text{sinc}\left(\frac{\pi(n_2 - 50)}{2}\right)$$

9. (5 points) Suppose that $\mathcal{X}$ is the unit disk, i.e.,

$$\mathcal{X} = \left\{ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1^2 + x_2^2 \leq 1 \right\}$$

Suppose that $\mathcal{Y}$ is defined as:

$$\mathcal{Y} = \left\{ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : \mathbf{y} = A\mathbf{x} \forall \mathbf{x} \in \mathcal{X} \right\}$$

where $\mathbf{A}$ is defined to be the following matrix:

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

Notice that the area of $\mathcal{X}$, in the two-dimensional plane, is $|\mathcal{X}| = \pi$. What is the numerical value of $|\mathcal{Y}|$, the area of $\mathcal{Y}$?

Solution:

$$|\mathcal{Y}| = |A|\pi = 3\pi$$

10. (5 points) Suppose that you are trying to allocate money to a set of N different possible investments. Suppose that if you allocate a_k dollars to investment k, it will return $a_k b_k$ dollars in profit. However, in order to invest, you first need to borrow the money, and the cost of borrowing is quadratic in terms of the allocations. Let $\mathbf{a}$ be your vector of allocations, let $\mathbf{b}$ be the vector of profit factors, and let C be the matrix of cost factors; suppose that your total profit is

$$P = \mathbf{b}^T \mathbf{a} - \mathbf{a}^T C \mathbf{a}$$

In terms of $\mathbf{b}$ and C, find the vector $\mathbf{a}$ that will maximize your profit. You may assume that C is nonsingular.

Solution:

$$\nabla_{\mathbf{a}} P = \mathbf{b} - 2C\mathbf{a}$$

Setting $\nabla_{\mathbf{a}} P = 0$ gives

$$\mathbf{a} = (\mathbf{C} + \mathbf{C}^T)^{-1} \mathbf{b}$$

If $\mathbf{C}$ is symmetric (a reasonable assumption, but not specified in the problem statement), then this is equivalent to

$$\mathbf{a} = \frac{1}{2} \mathbf{C}^{-1} \mathbf{b}$$

11. (24 points) The images $y[\eta]$ and $x[\mathbf{m}]$ are related by an affine transformation, where $\eta = [\eta, \xi, 1]^T$ and $\mathbf{m} = [m, n, 1]$ are coordinate vectors of the input and output image, respectively, m is the row index, and n is the column index.

(a) The affine transformation $\eta = \mathbf{Am}$ is a rotation by $-\frac{\pi}{3}$ radians. Find $\mathbf{A}$.

Solution:

$$\mathbf{A} = \begin{bmatrix} \cos(-\pi/3) & -\sin(-\pi/3) & 0 \\ \sin(-\pi/3) & \cos(-\pi/3) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(b) The affine transformation $\eta = \mathbf{Bm}$ consists of scaling the height of the image (m) by a factor of 5, while keeping the width (n) unchanged. Find $\mathbf{B}$.

Solution:

$$\mathbf{B} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(c) The affine transformation $\eta = \mathbf{Cm}$ consists of shifting all pixels to the left (negative n direction) by 20 columns. Find $\mathbf{C}$.

Solution:

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -20 \\ 0 & 0 & 1 \end{bmatrix}$$

(d) The affine transformation $\eta = D\mathbf{m}$ consists of performing parts (a) through (c) of this problem, one after the other, in order. Specify the matrix $\mathbf{D}$ in terms of the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. **There should be no numbers in your answer to this part**.

Solution:

$$\mathbf{D} = \mathbf{CBA}$$

12. (11 points) A particular triangle has corner coordinates at

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Let $\beta_0 = [\beta_1, \beta_2, \beta_3]^T$ be the barycentric coordinate vector corresponding to pixel $\mathbf{x}_0 = \left[\frac{2}{3}, \frac{1}{3}\right]^T$. Find β_0.

Solution:

$$\beta_0 = \begin{bmatrix} 0 \\ 1/3 \\ 2/3 \end{bmatrix}$$

13. (12 points) The images $y[\eta]$ and $x[\mathbf{m}]$ are related by an affine transformation $\eta = \mathbf{A}\mathbf{m}$, where $\eta = [\eta, \xi, 1]^T$ and $\mathbf{m} = [m, n, 1]^T$ are coordinate vectors of the input and output images, respectively. It is known that under this transformation, the origin swaps places with the point $[2, 2]$, thus

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \to \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 2 \\ 2 \end{bmatrix} \to \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Specify the $\mathbf{A}$ matrix as completely as you can. There should be two scalar variables in your answer; you may use the variables names α and β.

Solution:

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

14. (15 points) A triangle begins at

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Notice that, in this triangle, the barycentric coordinates of any point (x_4, y_4) are given by

$$\beta_4 = \begin{bmatrix} 1 - x_4 - y_4 \\ x_4 \\ y_4 \end{bmatrix}$$

Suppose the triangle is rotated, shifted and scaled to the new position

$$\mathbf{\Xi} = \begin{bmatrix} \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

The point $(x_4, y_4) = \left(\frac{1}{3}, \frac{1}{3}\right)$, internal to the first triangle, gets mapped to some point (ξ_4, η_4). Find ξ_4 and η_4.

Solution: $\xi_4 = \frac{1}{3}, \quad \eta_4 = 2$

15. (10 points) A reference image $I_0(u, v)$ has the following pixel values:

$$I_0(u, v) = 1 + (-1)^{u+v}$$

The test image $I_1(x, y)$ is created by piece-wise affine transformation of the pixel locations in $I_0(u, v)$. In particular, the triangle U in $I_0(u, v)$ is moved to the triangle X in $I_1(x, y)$, where

$$U = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 3 \\ 1 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}$$

(a) Find the reference coordinate $\mathbf{u} = [u, v, 1]^T$ that corresponds to the test coordinate $\mathbf{x} = [3, 2, 1]^T$.

Solution:

$$\beta^T = \left[\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right]$$
$$\mathbf{u}^T = [1.5, 2, 1]$$

(b) Use bilinear interpolation to find the value of the test pixel $I_1(3, 2)$.

Solution:

$$I_1(3, 2) = I_0(1.5, 2) = \frac{1}{2} I_0(1, 2) + \frac{1}{2} I_0(2, 2) = 1$$

16. (16 points) Remember that an affine transform is defined by a matrix with the following form:

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix}$$

Define the scalar term β to be $\beta = bd - (a-1)(e-1)$. It turns out that, as long as $\beta \neq 0$, there is exactly one input vector of the form $\mathbf{u}_0 = [u_0, v_0, 1]^T$ that maps to itself ($\mathbf{A}\mathbf{u}_0 = \mathbf{u}_0$). Find u_0 and v_0 in terms of a, b, c, d, e, f and β. HINT: you may find it useful to know that the inverse of a 2×2 matrix is

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}^{-1} = \frac{1}{\alpha\delta - \beta\gamma} \begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix}$$

Solution:

$$u_0 = \frac{(e-1)c - bf}{\beta}, \quad v_0 = \frac{(a-1)f - dc}{\beta}$$

17. (17 points) The barycentric coordinates of point $\mathbf{x}_0 = [x_0, y_0, 1]^T$, as defined by the triangle $\mathbf{x}_1 = [x_1, y_1, 1]^T$,$\mathbf{x}_2 = [x_2, y_2, 1]^T$,$\mathbf{x}_3 = [x_3, y_3, 1]^T$, are the coordinates β_1, β_2, β_3 such that $\mathbf{x}_0 = \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3$. If we constrain $\beta_1 + \beta_2 + \beta_3 = 1$, then there are actually only two degrees of freedom; for example, we could substitute $\beta_3 = 1-\beta_1-\beta_2$. A more interesting way to specify the two degrees of freedom is by defining variables a and b, $0 \le a, b \le 1$, such that

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} ab \\ (1-a)b \\ 1-b \end{bmatrix}$$

Draw a two-dimensional Cartesian plane, and label the x and y axes. Label the point $\mathbf{x}_1 = [0,0,1]^T$, $\mathbf{x}_2 = [2,0,1]^T$, $\mathbf{x}_3 = [1,2,1]^T$, and $\mathbf{x}_0 = [1,1,1]^T$. Now, given the values of these four points, find the values of a and b, and sketch the line segment connecting the point $\mathbf{x}_3$ to the point $a\mathbf{x}_1 + (1-a)\mathbf{x}_2$.

Solution:

$$a = \frac{1}{2}$$
$$b = \frac{1}{2}$$

The sketch should show a line segment connecting the point $\mathbf{x}_3 = [1,2]^T$ to the point $a\mathbf{x}_1 + (1-a)\mathbf{x}_2 = [1,0]^T$.

18. (17 points) The Barycentric coordinates of point $\mathbf{x}_0 = [x_0, y_0, 1]^T$, as defined by the triangle $\mathbf{x}_1 = [x_1, y_1, 1]^T$,$\mathbf{x}_2 = [x_2, y_2, 1]^T$,$\mathbf{x}_3 = [x_3, y_3, 1]^T$, are the coordinates β_1, β_2, β_3 such that

$$\begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$$

Provide an equation in terms of the six scalars $x_1, x_2, x_3, y_1, y_2, y_3$ specifying the conditions under which the matrix $\begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}$ is singular.

Solution: There are several such equations. One shows that the columns are linearly dependent:

$$\exists \alpha, \beta \text{ s.t. } \begin{matrix} x_3 = \alpha x_1 + \beta x_2 \\ y_3 = \alpha y_1 + \beta y_2 \\ 1 = \alpha + \beta \end{matrix}$$

Another shows that the determinant is zero:

$$x_1(y_2 - y_3) - x_2(y_1 - y_3) + x_3(y_1 - y_2) = 0$$

19. (16 points) Consider four points, $\mathbf{u}_1 = [u_1, v_1, 1]^T$, $\mathbf{u}_2 = [u_2, v_2, 1]^T$, $\mathbf{u}_3 = [u_1 + \alpha\cos\theta, v_1 + \alpha\sin\theta, 1]^T$, and $\mathbf{u}_4 = [u_2 + \beta\cos\theta, v_2 + \beta\sin\theta, 1]^T$. Notice that the slope of the line segment connecting $\mathbf{u_1}$ to $\mathbf{u}_3$ is $\frac{\alpha\sin\theta}{\alpha\cos\theta} = \tan\theta$, while the slope of the line segment connecting $\mathbf{u_2}$ to $\mathbf{u}_4$ is also $\frac{\beta\sin\theta}{\beta\cos\theta} = \tan\theta$. Suppose that there is an affine transform A such that $\mathbf{x}_1 = A\mathbf{u}_1$, $\mathbf{x}_2 = A\mathbf{u}_2$, $\mathbf{x}_3 = A\mathbf{u}_3$, and $\mathbf{x}_4 = A\mathbf{u}_4$. Prove that, for any affine transform matrix A, the line segment connecting $\mathbf{x}_1$ to $\mathbf{x}_3$ is parallel to (has the same slope as) the line segment that connects $\mathbf{x}_2$ to $\mathbf{x}_4$.

Solution: Use

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix}$$

Define $\mathbf{du}_3 = \mathbf{u}_3 - \mathbf{u}_1$, $\mathbf{du}_4 = \mathbf{u}_4 - \mathbf{u}_2$, $\mathbf{dx}_3 = \mathbf{x}_3 - \mathbf{x}_1 = A\mathbf{du}_3$, $\mathbf{dx}_4 = \mathbf{x}_4 - \mathbf{x}_2 = A\mathbf{du}_4$. Then

$$\mathbf{dx}_3 = \begin{bmatrix} a\alpha\cos\theta + b\alpha\sin\theta \\ d\alpha\cos\theta + e\alpha\sin\theta \\ 0 \end{bmatrix}, \quad \mathbf{dx}_4 = \begin{bmatrix} a\beta\cos\theta + b\beta\sin\theta \\ d\beta\cos\theta + e\beta\sin\theta \\ 0 \end{bmatrix}$$

Then the slopes are given by

$$\text{slope}(\bar{x_1 x_3}) = \frac{d\alpha\cos\theta + e\alpha\sin\theta}{a\alpha\cos\theta + b\alpha\sin\theta} = \frac{d\cos\theta + e\sin\theta}{a\cos\theta + b\sin\theta}$$
$$\text{slope}(\bar{x_2 x_4}) = \frac{d\beta\cos\theta + e\beta\sin\theta}{a\beta\cos\theta + b\beta\sin\theta} = \frac{d\cos\theta + e\sin\theta}{a\cos\theta + b\sin\theta}$$

20. (20 points) A particular signal, $x[n]$, is sampled at $F_s = 18,000$ samples/second. There are a total of $10,000$ samples, numbered $x[0]$ through $x[9999]$. These samples are divided into T frames, $\mathbf{x}_t$, with a window length of 250 samples and a hop length of 100 samples, i.e.,

$$\mathbf{x}_t = \begin{bmatrix} x[100t] \\ \vdots \\ x[100t+249] \end{bmatrix}$$

Your goal is to create two different matrices: $X = [\mathbf{X}_0, \ldots, \mathbf{X}_{T-1}]$ is the STFT (short-time Fourier transform) of $x[n]$, and $S = [\mathbf{S}_0, \ldots, \mathbf{S}_{T-1}]$ is the spectrogram of $x[n]$. The final image matrix S should show the spectral level (in decibels) of $x[n]$, as a function of time and frequency.

(a) Find T, the number of frames. This should be set so that (1) every sample of $x[n]$ appears in at least one frame, and (2) there is at most one frame with zero-padding. Your answer should be a number, or an explicit numerical expression.

Solution: The first frame is at $t = 0$, and the last is where $100t+249 \geq 9999$, i.e., $t \geq 97.5$. The smallest such integer is $t = 98$, therefore there are $T = 99$ frames.

(b) Your goal is to create 480-row matrices (length of $\mathbf{X}_t$ is 480), representing the spectrum between 0Hz and 5000Hz. Each STFT vector, $\mathbf{X}_t$, contains the frequency bins in the range between 0 and 5000Hz, from the length-N DFT of one frame of $\mathbf{x}_t$. Find N. Your answer should be a number, or an explicit numerical expression.

Solution:

$$N = 480\left(\frac{18000}{5000}\right) = 1728$$

(c) The STFT is given by $\mathbf{X}_t = \mathbf{A}\mathbf{x}_t$ for some matrix, A, whose $(k,n)^{\text{th}}$ element is a_{kn}. Give an expression for a_{kn} in terms of k, n, and N.

Solution:

$$a_{kn} = e^{-j\frac{2\pi kn}{N}}$$

(d) Suppose that $X_{max} = \max_k \max_t |X[k,t]|$. The spectrogram $S[k,t]$ is the level of $X[k,t]$, in decibels. Create an image $I[k,t]$ whose colors are scaled, shifted, and clipped versions of the pixels of $S[k,t]$ such that that the resulting image, $I[k,t]$, is equal to 255 if $X[k,t] = X_{max}$, and is equal to zero if $|X[k,t]| \leq X_{max}/1000$. Give an equation specifying $I[k,t]$ as a function of $X[k,t]$.

Solution:

$$I[k,t] = \max\left(0, 255\left(\frac{\log\left(\frac{|X[k,t]|}{X_{MAX}}\right)}{\log(1000)} + 1\right)\right)$$

21. (5 points) The signal $x[n]$ is given by

$$x[n] = \begin{cases} \cos(\omega_0 n) & 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases}$$

$X[k]$ is the length-N DFT of $x[n]$. Find $X[k]$, in terms of N and ω_0. You may find it useful to write your answer in terms of the transform of a rectangular window, $W_R(\omega)$, which is

$$w_R[n] = \begin{cases} 1 & 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases} \leftrightarrow W_R(\omega) = \frac{\sin(\omega N/2)}{\sin(\omega/2)} e^{-j\omega\left(\frac{N-1}{2}\right)}$$

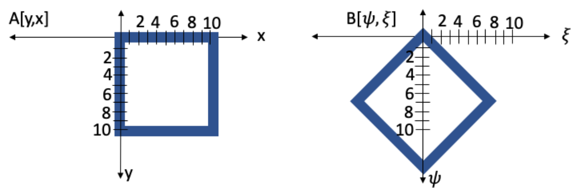
Solution:

$$\begin{aligned} X[k] &= \frac{1}{2} W_R\left(\frac{2\pi k}{N} - \omega_0\right) + \frac{1}{2} W_R\left(\frac{2\pi k}{N} + \omega_0\right) \\ &= \frac{1}{2} W_R\left(\frac{2\pi k}{N} - \omega_0\right) + \frac{1}{2} W_R\left(\frac{2\pi k}{N} - (2\pi - \omega_0)\right) \end{aligned}$$

22. (15 points) Suppose you have a picture of a white square on a black field, $A[y,x]$, where x is the column index, y is the row index. You wish to perform an affine transform that will turn your square into a picture of a diamond, $B[\psi,\xi]$, in which ξ is the column index, and ψ is the row index:

$$A[y,x] = \begin{cases} 255 & x=0 \text{ or } x=10,\quad 0\le y\le 10 \\ 255 & y=0 \text{ or } y=10,\quad 0\le x\le 10 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$B[\psi,\xi] = \begin{cases} 255 & \psi-\xi=0 \text{ or } \psi-\xi=10\sqrt{2},\quad 0\le \psi+\xi\le 10\sqrt{2} \\ 255 & \psi+\xi=0 \text{ or } \psi+\xi=10\sqrt{2},\quad 0\le \psi-\xi\le 10\sqrt{2} \\ 0 & \text{otherwise} \end{cases}$$



(a) **Affine Transform:** This affine transform can be written by a transform matrix, as

$$\begin{bmatrix} \xi \\ \psi \\ 1 \end{bmatrix} = \begin{bmatrix} a,b,c \\ d,e,f \\ g,h,i \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Find a,b,c,d,e,f,g,h and i. Your answers should be numbers, or else explicit numerical expressions; there should be no unresolved variables in your answers. Note: there is more than one correct answer.

Solution: There are exactly eight correct solutions. The two solutions that map point $x=0, y=0$ to point $\xi=0, \psi=0$ are given by

$$\begin{bmatrix} a,b,c \\ d,e,f \\ g,h,i \end{bmatrix} = \begin{bmatrix} \pm\cos(\pi/4), \mp\sin(\pi/4), 0 \\ \sin(\pi/4), \cos(\pi/4), 0 \\ 0,0,1 \end{bmatrix}$$

There are also six more solutions: 3 different corners of the original square can map to the point (0,0), each with 2 different corners mapped to the point at $(-5\sqrt{2}, 5\sqrt{2})$.

(b) **Bilinear Interpolation:** $A[y,x]$ is a discrete-space image (y and x are integers), whereas $A(y,x)$ is the corresponding continuous-space image (y and x are real numbers). An affine transform maps integer coordinates ξ and ψ to real-valued coordinates x and y, so it's necessary to estimate the values of the continuous-space image. For example, one way to compute the intensity of the pixel at $B[2,1]$ is by setting it equal to $A\left(\frac{3\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)\approx A\left(2.1,0.7\right)$. Use bilinear interpolation, combined with the information given in Eq. 1, to estimate the numerical value of $A\left(2.1,0.7\right)$.

> **Solution:**
>
> $$\begin{aligned}A(2.1,0.7) &= (0.9)(0.3)A[2,0]+(0.9)(0.7)A[2,1]+(0.1)(0.3)A[3,0]+(0.1)(0.7)A[3,1]\\ &= (0.9)(0.3)(255)+(0.9)(0.7)(0)+(0.1)(0.3)(255)+(0.1)(0.7)(0)\\ &= (0.3)255\end{aligned}$$
>
> (The last step, simplification, is optional).

(c) **Barycentric Coordinates:** Suppose we have some coordinate with known values of x and y, and we're trying to find the values of ξ and ψ to which it gets moved. One way to solve this problem is by using the transform matrix you computed in part (a). A different solution is to use the triangle whose coordinates are $[x_1,y_1]$, $[x_2,y_2]$, and $[x_3,y_3]$ before transformation, but $[\xi_1,\psi_1]$, $[\xi_2,\psi_2]$, and $[\xi_3,\psi_3]$ after transformation, where

$$\begin{bmatrix}x_1\\y_1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix},\quad \begin{bmatrix}x_2\\y_2\end{bmatrix}=\begin{bmatrix}10\\0\end{bmatrix},\quad \begin{bmatrix}x_3\\y_3\end{bmatrix}=\begin{bmatrix}0\\10\end{bmatrix}$$

$$\begin{bmatrix}\xi_1\\\psi_1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix},\quad \begin{bmatrix}\xi_2\\\psi_2\end{bmatrix}=\begin{bmatrix}5\sqrt{2}\\5\sqrt{2}\end{bmatrix},\quad \begin{bmatrix}\xi_3\\\psi_3\end{bmatrix}=\begin{bmatrix}-5\sqrt{2}\\5\sqrt{2}\end{bmatrix}$$

In terms of these triangles, the Barycentric coordinates of any point are the same before and after transformation:

$$\begin{bmatrix}x\\y\end{bmatrix}=\beta_1\begin{bmatrix}x_1\\y_1\end{bmatrix}+\beta_2\begin{bmatrix}x_2\\y_2\end{bmatrix}+\beta_3\begin{bmatrix}x_3\\y_3\end{bmatrix},\quad \begin{bmatrix}\xi\\\psi\end{bmatrix}=\beta_1\begin{bmatrix}\xi_1\\\psi_1\end{bmatrix}+\beta_2\begin{bmatrix}\xi_2\\\psi_2\end{bmatrix}+\beta_3\begin{bmatrix}\xi_3\\\psi_3\end{bmatrix}$$

where $\beta_1+\beta_2+\beta_3=1$. Suppose that x and y are known, but ξ and ψ are unknown. **Find β_1, β_2, and β_3 in terms of x and y.**

> **Solution:**
>
> $$\beta_2=\frac{x}{10},\quad \beta_3=\frac{y}{10},\quad \beta_1=1-\beta_2-\beta_3$$

23. (25 points) Consider applying the Griffin-Lim algorithm to reconstruct $x[n]$ from $X_m[k]$, an STFT with window length of 2 samples and hop length of 1 sample. Suppose that we start with the following initial estimate of $X_m[k]$:

$$X_m[0] = \begin{cases} 2 & m \text{ even} \\ 1 & m \text{ odd} \end{cases}, \quad X_m[1] = \begin{cases} 1 & m \text{ even} \\ 2 & m \text{ odd} \end{cases}$$

Not that the initial phase is zero at all frequencies; don't try to apply a random initial phase.

(a) The first step in Griffin-Lim is to find

$$x[n] = \text{ISTFT}\left(X_m[k]\right)$$

Compute $x[n]$ for all $n \geq 1$. Assume that the windows are rectangular, and that therefore the STFT denominator is $\sum_m w[n-m] = 2$.

Solution: For even values of m, $X_m[k] = [2, 1]$, which has the inverse DFT $x[n+m] = \left[\frac{3}{2}, \frac{1}{2}\right]$. For odd values of m, $X_m[k] = [1, 2]$, which has the inverse DFT $x[n+m] = \left[\frac{3}{2}, -\frac{1}{2}\right]$. Overlapping these signals by one sample, adding them together, and dividing by two, we get:

$$x[n] = \begin{cases} \frac{1}{2}\left(\frac{3}{2} - \frac{1}{2}\right) = \frac{1}{2} & n \text{ even} \\ \frac{1}{2}\left(\frac{3}{2} + \frac{1}{2}\right) = 1 & n \text{ odd} \end{cases}$$

(b) The second step in Griffin-Lim is to find

$$\tilde{X}_m[k] = \text{STFT}\,(x[n])$$

Find $\tilde{X}_m[k]$ for $m \geq 1$. Use rectangular windows.

Solution:

$$\begin{aligned} X_m[k] &= \text{DFT}\,(x[n+m]w[n]) \\ &= \begin{cases} \text{DFT}\left([1, \frac{1}{2}]\right) = \left[\frac{3}{2}, \frac{1}{2}\right] & m \text{ odd} \\ \text{DFT}\left([\frac{1}{2}, 1]\right) = \left[\frac{3}{2}, -\frac{1}{2}\right] & m \text{ even} \end{cases} \end{aligned}$$