Problem Set 6
Signals

For the Exercise Session on Nov 26 — **Due: Tue, Dec 2, 10am, on Moodle**

1 Problem for Class

Problem 1: The Fourier matrix diagonalizes all circulant matrices.

The discrete Fourier transform (DFT) $\mathbf{X}$ of the vector $\mathbf{x}$ is given by

$$\mathbf{X} = W\mathbf{x} \quad \text{and} \quad \mathbf{x} = \frac{1}{N} W^H \mathbf{X}. \tag{1}$$

In this homework problem, you will prove that the Fourier matrix diagonalizes all *circulant* matrices.

(a) To cut the derivation into two simpler steps, we introduce an auxiliary matrix M, defined as

$$M = WA \quad = \quad W \underbrace{\begin{pmatrix} b_0 & b_{N-1} & b_{N-2} & b_{N-3} & \dots & b_1 \\ b_1 & b_0 & b_{N-1} & b_{N-2} & \dots & b_2 \\ b_2 & b_1 & b_0 & b_{N-1} & \dots & b_3 \\ b_3 & b_2 & b_1 & b_0 & \dots & b_4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{N-1} & b_{N-2} & b_{N-3} & b_{N-4} & \dots & b_0 \end{pmatrix}}_{\text{This is a circulant matrix}}. \tag{2}$$

Let us denote the unitary DFT of the sequence $\{b_0, b_1, \dots, b_{N-1}\}$ by $\{B_0, B_1, \dots, B_{N-1}\}$. Write out the matrix M in terms of $\{B_0, B_1, \dots, B_{N-1}\}$. *Hint:* The first column of the matrix M is simply given by

$$W \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{N-1} \end{pmatrix} = \begin{pmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_{N-1} \end{pmatrix} \tag{3}$$

To find the second column, you will need to use some Fourier properties.

(b) Using the matrix M from above, compute the full matrix product

$$WAW^H \quad = \quad MW^H. \tag{4}$$

Hint: Handle every *row* of the matrix M separately. Define the vector $\mathbf{m}$ such that $\mathbf{m}^H$ is simply the first row of the matrix M. But the product $\mathbf{m}^H W^H$ is easily computed, recalling that $\mathbf{m}^H W^H = (W\mathbf{m})^H$.

Solution 1. *(a)* Note that $\forall i \in \{0, \dots, N-1\}$:

$$B_i = \sum_{j=0}^{N-1} W_{i,j} b_j \tag{5}$$

Let the entry of W in i^{th} row, j^{th} column denoted by $W_{i,j} = w^{ij}$, where $w = e^{-j\frac{2\pi}{N}}$. For the second column of M, each b_j is multiplied with $W_{i,j+1}$ instead of $W_{i,j}$. Also, $W_{i,j+1} = W_{i,j} w^i$. Using the cyclic shift property, we find

$$M = \begin{pmatrix} B_0 & w^0 B_0 & w^{0\times 2} B_0 & \cdots & w^{0\times(N-1)} B_0 \\ B_1 & w^1 B_1 & w^{1\times 2} B_1 & \cdots & w^{1\times(N-1)} B_1 \\ B_2 & w^2 B_2 & w^{2\times 2} B_2 & \cdots & w^{2\times(N-1)} B_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{N-1} & w^{N-1} B_{N-1} & w^{(N-1)\times 2} B_{N-1} & \cdots & w^{(N-1)\times(N-1)} B_{N-1} \end{pmatrix} \tag{6}$$

$$= \begin{pmatrix} B_0 & B_0 & B_0 & \cdots & B_0 \\ B_1 & w^1 B_1 & w^2 B_1 & \cdots & w^{N-1} B_1 \\ B_2 & w^2 B_2 & w^4 B_2 & \cdots & w^{2(N-1)} B_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ B_{N-1} & w^{N-1} B_{N-1} & w^{2(N-1)} B_{N-1} & \cdots & w^{(N-1)(N-1)} B_{N-1} \end{pmatrix} \tag{7}$$

(b) Using the hint, let us tackle the first row, thus $\mathbf{m}^H = (B_0, B_0, \dots, B_0)$. We need to find $W\mathbf{m}$, which is simply the DFT of the vector $\mathbf{m}$.

$$W\mathbf{m} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & w & \dots & w^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & w^{N-1} & \dots & w^{(N-1)(N-1)} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} B_0 \tag{8}$$

$$= B_0 \begin{pmatrix} \sum_{i=0}^{N-1} w^0 \\ \sum_{i=0}^{N-1} w^i \\ \vdots \\ \sum_{i=0}^{N-1} w^{(N-1)i} \end{pmatrix} \tag{9}$$

$$= \begin{pmatrix} NB_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{10}$$

This is the vector $(NB_0, 0, 0, \dots, 0)$, which is thus the first row of the product MW^H. Note that the j^{th} row vector of M can be expressed as $\mathbf{m}_j^H = (w^{(j-1)\times 0} B_j, w^{(j-1)\times 1} B_j, \dots, w^{(j-1)\times(N-1)} B_j)$. Hence we have

$$Wm_j = B_j \begin{pmatrix} \sum_{i=0}^{N-1} w^{(0+(j-1))\times i} \\ \sum_{i=0}^{N-1} w^{(1+(j-1))\times i} \\ \vdots \\ \sum_{i=0}^{N-1} w^{(N-1+(j-1))\times i} \end{pmatrix} \tag{11}$$

Since for each $j \in \{1, 2, \dots, N\}$, it can be verified that only the j^{th} summation is NB_j and others are 0. Therefore, we have

$$WAW^H = MW^H = \begin{pmatrix} NB_0 & 0 & \dots & 0 \\ 0 & NB_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & NB_{N-1} \end{pmatrix} \tag{12}$$

2 The Homework

Problem 2:

(a) What is the parametric-form maximum entropy density $f(x)$ satisfying the two conditions

$$\mathbb{E}[X^8] = a \qquad \mathbb{E}[X^{16}] = b$$

(b) What is the maximum entropy density satisfying the condition

$$\mathbb{E}[X^8 + X^{16}] = a + b$$

(c) Which entropy is higher?

Solution 2. *(a)* The maximum entropy distribution subject to $\mathbb{E}[X^8] = a$ and $\mathbb{E}[X^{16}] = b$ has the parametric-form

$$f(x) \quad = e^{\theta_1 x^8 + \theta_2 x^{16} - A(\theta_1, \theta_2)}. \tag{13}$$

Note that here, it is not possible to give a closed-form formula for the log-partition function

$$A(\theta_1, \theta_2) = \ln \int_{-\infty}^{\infty} e^{\theta_1 x^8 + \theta_2 x^{16}} dx$$

simply because this integral does not have a closed-form solution. The last step would be to select θ_1 and θ_2 so that

$$\int_{-\infty}^{\infty} x^8 e^{\theta_1 x^8 + \theta_2 x^{16} - A(\theta_1, \theta_2)} dx = a,$$
$$\int_{-\infty}^{\infty} x^{16} e^{\theta_1 x^8 + \theta_2 x^{16} - A(\theta_1, \theta_2)} dx = b.$$

In principle, this is all good, but we cannot solve these integrals in closed form, either. So, we'd have to resort to numerical methods.

(b) The maximum entropy distribution subject to $\mathbb{E}[X^8 + X^{16}] = a + b$ has the parametric-form

$$f(x) \quad = e^{\theta(x^8 + x^{16}) - A(\theta)} \tag{14}$$

Note that again, it is not possible to give a closed-form formula for the log-partition function

$$A(\theta) = \ln \int_{-\infty}^{\infty} e^{\theta(x^8 + x^{16})} dx.$$

(c) The resulting maximum entropy subject to the condition in (b) is higher. This is because any distribution satisfying the conditions in (a) always also satisfies the condition in (b). (But note that we cannot give a nice closed-form formula for the two entropies since we cannot explicitly calculate the (log-)partition function.

3 Additional Problems

Problem 3: Exponential Families and Maximum Entropy

Let $Y = X_1 + X_2$. Find the maximum entropy of Y under the constraint $\mathbb{E}[X_1^2] = P_1$, $\mathbb{E}[X_2^2] = P_2$:

(a) If X_1 and X_2 are independent.

(b) If X_1 and X_2 are allowed to be dependent.

Solution 3. *(a)* If X_1 and X_2 are independent,

$$\mathrm{Var}[Y] = \mathrm{Var}[X_1 + X_2] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] \leq \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] = P_1 + P_2 \tag{15}$$

where equality holds when $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$. Thus we have

$$\max_{f(y)} h(Y) \quad \leq \quad \frac{1}{2}\log(2\pi e(P_1 + P_2)) \tag{16}$$

where equality holds when Y is Gaussian with zero mean, which requires X_1 and X_2 to be independent and Gaussian with zeros mean.

(b) For dependent X_1 and X_2, we have

$$\mathrm{Var}(Y) \leq \mathbb{E}[Y^2] = \mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] + 2\mathbb{E}[X_1 X_2] \leq (\sqrt{P_1} + \sqrt{P_2})^2 \tag{17}$$

where the first equality holds when $\mathbb{E}[Y] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = 0$, and the send equality holds when $X_2 = \sqrt{\frac{P_2}{P_1}} X_1$. Hence, $\max_{f(y)} h(Y) \leq \frac{1}{2}\log(2\pi e(\sqrt{P_1} + \sqrt{P_2})^2)$, where equality holds when Y is Gaussian with zero mean, which requires X_1 and X_2 to be Gaussian with zero mean and $X_2 = \sqrt{\frac{P_2}{P_1}} X_1$.

Problem 4: Exponential Families and Maximum Entropy

For $t > 0$, consider a family of distributions supported on $[t, +\infty]$ such that $\mathbb{E}[\ln X] = \frac{1}{\alpha} + \ln t$, $\alpha > 0$.

1. What is the parametric form of a maximum entropy distribution satisfying the constraint on the support and the mean?
2. Find the exact form of the distribution.

Solution 4. (i) The maximum entropy distribution has the parametric form $e^{\theta \ln x - A(\theta)} = x^\theta e^{-A(\theta)}$.

(ii) Let us first find the value of $A(\theta)$ from the density constraint $\int_t^\infty x^\theta e^{-A(\theta)} dx = 1$. This gives $e^{-A(\theta)} = -\frac{\theta+1}{t^{\theta+1}}$.

Next we find θ from the mean constraint $\int_t^\infty x^\theta e^{-A(\theta)} \ln x \, dx = \frac{1}{\alpha} + \ln t$. This gives $\frac{t^{\theta+1}((\theta+1)\ln t - 1)}{t^{\theta+1}(\theta+1)} = \ln t - \frac{1}{\theta+1} = \frac{1}{\alpha} + \ln t$ and therefore $\theta = -(\alpha+1)$. The resulting form of the distribution is

$$p(x) = \frac{\alpha t^\alpha}{x^{\alpha+1}}$$

.

Problem 5: Minimum-norm Solutions

In this problem, we consider an *underdetermined* system of linear equations, i.e., $A\mathbf{x} = \mathbf{b}$, where $A_{m\times n}$ is a "wide" matrix ($m < n$) and $\mathbf{b}$ is chosen such that a solution exists. As you know, in this case, there exist infinitely many solutions. Prove that the one solution $\mathbf{x}$ that has the minimum 2-norm can be expressed as

$$\mathbf{x}_{MN} \quad = \quad V\Sigma^{-1}U^H\mathbf{b}, \tag{18}$$

where, as usual, the SVD of $A = U\Sigma V^H$, and Σ^{-1} is the matrix Σ where all non-zero diagonal entries are inverted.

Hint: Clearly, A is not a full-rank matrix, and thus cannot be inverted. However, it might be possible to *construct* a matrix A' such that $A'\mathbf{x} = \mathbf{b}'$ has a solution, A is a submatrix of A' and $\mathbf{b}$ is a subvector of $\mathbf{b}'$. What will be the norm of $\mathbf{x}$ in such a case?

Solution 5. Consider the singular value decomposition (SVD) $A = U\Sigma V^H$. Here, U is an $m \times m$ unitary matrix and Σ is given by

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_m \end{bmatrix}. \tag{19}$$

We assume that A is full rank, meaning that $\sigma_i > 0$ for all i. Finally, V is a matrix of dimension $n \times m$ whose columns are orthonormal, i.e., $V^H V = I_m$.

Note that $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions. Our goal is to find the one solution with the smallest 2-norm.

We will now construct a new matrix which is of dimension $n \times n$ and of full rank. To this end, pick any unitary matrix U_B of dimension $(n-m) \times (n-m)$ and any diagonal matrix Σ_B, also of dimension $(n-m)\times(n-m)$, with strictly positive entries on the diagonal. Moreover, pick a matrix V_B of dimension $n \times (n-m)$ whose columns are orthonormal and at the same time orthogonal to all of the columns in V. (It is easy to see that such a matrix exists.) Define the matrix $B = U_B \Sigma_B V_B^H$. With this, we can now stack up A and B, and they satisfy

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & U_B \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_B \end{bmatrix} \begin{bmatrix} V^H \\ V_B^H \end{bmatrix}, \tag{20}$$

where we also note that the last expression is indeed the singular value decomposition of the stacked matrix. Select any vector $\mathbf{b}_B$ of length $n-m$. Now consider solutions $\mathbf{x}$ to the system

$$\begin{bmatrix} A \\ B \end{bmatrix} \mathbf{x} = \begin{bmatrix} U & 0 \\ 0 & U_B \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_B \end{bmatrix} \begin{bmatrix} V^H \\ V_B^H \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_B \end{bmatrix}. \tag{21}$$

Since the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ is full rank by construction, thus invertible, the solution is given by

$$\mathbf{x} = [V, V_B] \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_B \end{bmatrix}^{-1} \begin{bmatrix} U & 0 \\ 0 & U_B \end{bmatrix}^H \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_B \end{bmatrix}. \tag{22}$$

Moreover, note that this solution $\mathbf{x}$ also satisfies $A\mathbf{x} = \mathbf{b}$, so it is a solution to our original system of equations. The square of the 2-norm of $\mathbf{x}$ is

$$\begin{aligned} ||\mathbf{x}||_2^2 = \mathbf{x}^H\mathbf{x} &= \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_B \end{bmatrix}^H \begin{bmatrix} U & 0 \\ 0 & U_B \end{bmatrix} \begin{bmatrix} \Sigma^H & 0 \\ 0 & \Sigma_B^H \end{bmatrix}^{-1} \begin{bmatrix} V^H \\ V_B^H \end{bmatrix} [V, V_B] \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_B \end{bmatrix}^{-1} \begin{bmatrix} U & 0 \\ 0 & U_B \end{bmatrix}^H \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_B \end{bmatrix} \\ &= ||V\Sigma^{-1}U^H\mathbf{b}||_2^2 + ||V_B\Sigma_B^{-1}U_B^H\mathbf{b}_B||_2^2 \end{aligned} \tag{23}$$

Both summands in the last expression are non-negative. Since the first summand is fixed, the expression is minimized if we can make the second summand zero. To do so, we select $\mathbf{b}_B = \mathbf{0}$. Hence in such case,

$$\mathbf{x}_{MN} = \begin{bmatrix} V \\ V_B \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma_B \end{bmatrix}^{-1} \begin{bmatrix} U & 0 \\ 0 & U_B \end{bmatrix}^H \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} = V\Sigma^{-1}U^H\mathbf{b}. \tag{24}$$

Problem 6: Exponential Families and Maximum Entropy 2

Find the maximum entropy density f, defined for $x \geq 0$, satisfying $\mathbb{E}[X] = \alpha_1$, $\mathbb{E}[\ln X] = \alpha_2$. That is, maximize $-\int f \ln f$ subject to $\int x f(x) dx = \alpha_1$, $\int (\ln x) f(x) dx = \alpha_2$, where the integral is over $0 \leq x < \infty$. What family of densities is this?

Solution 6. The maximum entropy distribution subject to constraints

$$\int xf(x)dx = \alpha_1 \tag{25}$$

and

$$\int (\ln x)f(x)dx = \alpha_2 \tag{26}$$

is of the form

$$f(x) = e^{\lambda_0+\lambda_1 x+\lambda_2 \ln x} = cx^{\lambda_2}e^{\lambda_1 x} \tag{27}$$

which is of the form of a Gamma distribution. The constants should be chosen so as to satisfy the constraints. We need to solve the following equations

$$\int_0^\infty f(x)dx = \int_0^\infty cx^{\lambda_2}e^{\lambda_1 x}dx = 1 \tag{28}$$

$$\int_0^\infty xf(x)dx = \int_0^\infty cx^{\lambda_2+1}e^{\lambda_1 x}dx = \alpha_1 \tag{29}$$

$$\int_0^\infty (\ln x)f(x)dx = \int_0^\infty cx^{\lambda_2}e^{\lambda_1 x}\ln x dx = \alpha_2 \tag{30}$$

Thus, the Gamma distributions $f(x) = \frac{1}{\Gamma(k)\theta^k}x^{k-1}e^{-\frac{x}{\theta}}$ with

$$\mathbb{E}[X] = k\theta = \alpha_1 \qquad \mathbb{E}[\ln X] = \psi(k) + \ln(\theta) = \alpha_2 \tag{31}$$

is the exponential family we want.