Appendix A

Glossary of Symbols and Abbrevations

Symbols

Boldface characters denote vectors or matrices. All others are scalars. All vectors are column vectors. Random variables are denoted by capital letters such as U, V, W, X, Y, Z and random vectors by $\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ and their values by corresponding lowercase letters.

$\angle$	angle of
$*$	complex conjugate
$\star$	convolution operator, either convolution sum or integral
$\hat{}$	denotes estimator
$\sim$	denotes *is distributed according to*
$[x]$	denotes the largest integer $\leq x$
x^+	denotes a number slightly larger than x
x^-	denotes a number slightly smaller than x
$A \times B$	cartesian product of sets A and B
$[\mathbf{A}]_{ij}$	(i,j)th element of $\mathbf{A}$
$\mathcal{A}(z)$	z-transform of $a[n]$ sequence
$[\mathbf{b}]_i$	ith element of $\mathbf{b}$
$\mathrm{Ber}(p)$	Bernoulli random variable
$\mathrm{bin}(M,p)$	binomial random variable
χ^2_N	chi-squared distribution with N degrees of freedom
$\binom{N}{k}$	number of combinations of N things taken k at a time
c	complement of set
$\mathrm{cov}(X,Y)$	covariance of X and Y
$\mathbf{C}$	covariance matrix

$\mathbf{C}_X$	covariance matrix of $\mathbf{X}$
$\mathbf{C}_{X,Y}$	covariance matrix of X and Y
$c_X[n_1, n_2]$	covariance sequence of discrete-time random process $X[n]$
$c_X(t_1, t_2)$	covariance function of continuous-time random process $X(t)$
$\delta(t)$	Dirac delta function or impulse function
$\delta[n]$	discrete-time unit impulse sequence
δ_{ij}	Kronecker delta
Δf	small interval in frequency f
Δt	small interval in t
Δx	small interval in x
Δ_t	time interval between samples
$\det(\mathbf{A})$	determinant of matrix $\mathbf{A}$
$\text{diag}(a_{11}, \ldots, a_{NN})$	diagonal matrix with elements a_{ii} on main diagonal
$\mathbf{e}_i$	natural unit vector in ith direction
η	signal-to-noise ratio
$E[\cdot]$	expected value
$E[X^n]$	nth moment
$E[(X - E[X])^n]$	nth central moment
$E_X[\cdot]$	expected value with respect to PMF or PDF of X
$E_{X,Y}[\cdot]$	expected value with respect to joint PMF or joint PDF of (X, Y)
$E_{X_1,X_2,\ldots,X_N}[\cdot]$	expected value with respect to N-dimensional joint PMF or PDF
$E_{\mathbf{X}}[\cdot]$	shortened notation for $E_{X_1,X_2,\ldots,X_N}[\cdot]$
$E_{Y\|X}[Y\|X]$	conditional expected value considered as random variable
$E_{Y\|X}[Y\|x_i]$	expected value of PMF $p_{Y\|X}[y_j\|x_i]$
$E_{Y\|X}[Y\|x]$	expected value of PDF $p_{Y\|X}(y\|x)$
$E[\mathbf{X}]$	expected value of random vector $\mathbf{X}$
$\in$	element of set
$\exp(\lambda)$	exponential random variable
f	discrete-time frequency
F	continuous-time frequency
$F_X(x)$	cumulative distribution function of X
$F_X^{-1}(x)$	inverse cumulative distribution function of X
$F_{X,Y}(x, y)$	cumulative distribution function of X and Y
$F_{X_1,\ldots,X_N}(x_1, \ldots, x_N)$	cumulative distribution function of $X_1, \ldots, X_N$
$F_{Y\|X}(y\|x)$	cumulative distribution function of Y conditioned on $X = x$
$\mathcal{F}$	Fourier transform
$\mathcal{F}^{-1}$	inverse Fourier transform
$g(\cdot)$	general notation for function of real variable
$g^{-1}(\cdot)$	general notation for inverse function of $g(\cdot)$

$\Gamma(x)$	Gamma function
$\Gamma(\alpha, \lambda)$	Gamma random variable
$\gamma_{X,Y}(f)$	coherence function for discrete-time random processes $X[n]$ and $Y[n]$
geom(p)	geometric random variable
$h[n]$	impulse response of LSI system
$h(t)$	impulse response of LTI system
$H(f)$	frequency response of LSI system
$H(F)$	frequency response of LTI system
$\mathcal{H}(z)$	system function of LSI system
$I_A(x)$	indicator function for the set A
$\mathbf{I}$	identity matrix
$\cap$	intersection of sets
j	$\sqrt{-1}$
$\frac{\partial(w,z)}{\partial(x,y)}$	Jacobian matrix of transformation of $w = g(x, y), z = h(x, y)$
$\frac{\partial(x_1,\ldots,x_N)}{\partial(y_1,\ldots,y_N)}$	Jacobian matrix of transformation from $\mathbf{y}$ to $\mathbf{x}$
$\mathbf{\Lambda}$	diagonal matrix with eigenvalues on main diagonal
mse	mean square error
μ	mean
$\mu_X[n]$	mean sequence of discrete-time random process $X[n]$
$\mu_X(t)$	mean function of continuous-time random process $X(t)$
$\boldsymbol{\mu}$	mean vector
$\binom{M}{k_1,k_2,\ldots,k_N}$	multinomial coefficient
n	discrete-time index
$N!$	N factorial
$(N)_r$	equal to $N(N-1)\cdots(N-r+1)$
N_A	number of elements in set A
$\mathcal{N}(\mu, \sigma^2)$	normal or Gaussian random variable with mean μ and variance σ^2
$\mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$	multivariate normal or Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance $\mathbf{C}$
$\|\|\mathbf{x}\|\|$	Euclidean norm or length of vector $\mathbf{x}$
$\emptyset$	null or empty set
opt	optimal value
$\mathbf{1}$	vector of all ones
Pois(λ)	Poisson random variable
$p_X[x_i]$	PMF of X
$p_X[k]$	PMF of integer-valued random variable X (or $p_X[i]$, $p_X[j]$)
$p_{X,Y}[x_i, y_j]$	joint PMF of X and Y
$p_{X_1,\ldots,X_N}[x_1, \ldots, x_N]$	joint PMF of $X_1, \ldots, X_N$
$p_{\mathbf{X}}[\mathbf{x}]$	shortened notation for $p_{X_1,\ldots,X_N}[x_1, \ldots, x_N]$
$p_{X_1,\ldots,X_N}[k_1, \ldots, k_N]$	joint PMF of integer-valued random variables $X_1, \ldots, X_N$

$p_{Y\|X}[y_j\|x_i]$	conditional PMF of Y given $X = x_i$
$p_{X_N\|X_1,\ldots,X_{N-1}}[x_N\| x_1,\ldots,x_{N-1}]$	conditional PMF of X_N given $X_1,\ldots,X_{N-1}$
$p_{X,Y}[i,j]$	joint PMF of integer-valued random variables X and Y
$p_{Y\|X}[j\|i]$	conditional PMF of integer-valued random variable Y given $X = i$
$p_X(x)$	PDF of X
$p_{X,Y}(x,y)$	joint PDF of X and Y
$p_{X_1,\ldots,X_N}(x_1,\ldots,x_N)$	joint PDF of $X_1,\ldots,X_N$
$p_{\mathbf{X}}(\mathbf{x})$	shortened notation for $p_{X_1,\ldots,X_N}(x_1,\ldots,x_N)$
$p_{Y\|X}(y\|x)$	conditional PDF of Y given $X = x$
$P[E]$	probability of the event E
P_e	probability of error
$P_X(f)$	power spectral density of discrete-time random process $X[n]$
$\mathcal{P}_X(z)$	z-transform of autocorrelation sequence $r_X[k]$
$P_X(F)$	power spectral density of continuous-time random process $X(t)$
$P_{X,Y}(f)$	cross-power spectral density of discrete-time random processes $X[n]$ and $Y[n]$
$P_{X,Y}(F)$	cross-power spectral density of continuous-time random processes $X(t)$ and $Y(t)$
$\phi_X(\omega)$	characteristic function of X
$\phi_{X,Y}(\omega_X,\omega_Y)$	joint characteristic function of X and Y
$\phi_{X_1,\ldots,X_N}(\omega_1,\ldots,\omega_N)$	joint characteristic function of $X_1,\ldots,X_N$
$\Phi(x)$	cumulative distribution function of $\mathcal{N}(0,1)$ random variable
$Q(x)$	probability that a $\mathcal{N}(0,1)$ random variable exceeds x
$Q^{-1}(u)$	value of $\mathcal{N}(0,1)$ random variable that is exceeded with probability of u
$\rho_{X,Y}$	correlation coefficient of X and Y
$r_X[k]$	autocorrelation sequence of discrete-time random process $X[n]$
$r_X(\tau)$	autocorrelation function of continuous-time random process $X(t)$
$r_{X,Y}[k]$	cross-correlation sequence of discrete-time random processes $X[n]$ and $Y[n]$
$r_{X,Y}(\tau)$	cross-correlation function of continuous-time random processes $X(t)$ and $Y(t)$
R or R^1	denotes real line
R^N	denotes N-dimensional Euclidean space
$\mathbf{R}_X$	autocorrelation matrix
$\mathcal{S}$	sample space

$\mathcal{S}_X$	sample space of random variable X
$\mathcal{S}_{X,Y}$	sample space of random variables X and Y
$\mathcal{S}_{X_1,X_2,\ldots,X_N}$	sample space of random variables $X_1, X_2, \ldots, X_N$
s_i	element of discrete sample space
s	element of continuous sample space
σ^2	variance
σ_X^2	variance of random variable X
$\sigma_X^2[n]$	variance sequence of discrete-time random process $X[n]$
$\sigma_X^2(t)$	variance function of continuous-time random process $X(t)$
$s[n]$	discrete-time signal
$\mathbf{s}$	vector of signal samples
$s(t)$	continuous-time signal
t	continuous time
^{T}	transpose of matrix
$\mathcal{U}(a,b)$	uniform random variable over the interval (a,b)
$\cup$	union of sets
$u[n]$	discrete unit step function
$u(x)$	unit step function
$\mathcal{U}(z)$	z-transform of $u[n]$ sequence
$\mathbf{V}$	modal matrix
$\text{var}(X)$	variance of X
$\text{var}(Y\|x_i)$	variance of conditional PMF or of $p_{Y\|X}[y_j\|x_i]$
x_i	value of discrete random variable
x	value of continuous random variable
X_s	standardized version of random variable X
x_s	value for X_s
$X[n]$	discrete-time random process
$x[n]$	realization of discrete-time random process
$X(t)$	continuous-time random process
$x(t)$	realization of continuous-time random process
$\mathcal{X}(z)$	z-transform of $x[n]$ sequence
$\bar{X}$	sample mean random variable
$\bar{x}$	value of $\bar{X}$
$\mathbf{X}$	random vector $(X_1, X_2, \ldots, X_N)$
$\mathbf{x}$	value $(x_1, x_2, \ldots, x_N)$ of random vector $\mathbf{X}$
$Y\|(X = x_i)$	random variable Y conditioned on $X = x_i$
$\mathcal{Z}$	z-transform
$\mathcal{Z}^{-1}$	inverse z-transform
$\mathbf{0}$	vector or matrix of all zeros

Abbreviations

ACF	autocorrelation function
ACS	autocorrelation sequence
AR	autoregressive
AR(p)	autoregressive process of order p
ARMA	autoregressive moving average
CCF	cross-correlation function
CCS	cross-correlation sequence
CDF	cumulative distribution function
CPSD	cross-power spectral density
CTCV	continuous-time/continuous-valued
CTDV	continuous-time/discrete-valued
D/A	digital-to-analog
dB	decibel
DC	constant level (direct current)
DFT	discrete Fourier transform
DTCV	discrete-time/continuous-valued
DTDV	discrete-time/discrete-valued
FFT	fast Fourier transform
FIR	finite impulse response
GHz	giga-hertz
Hz	hertz
IID	independent and identically distributed
IIR	infinite impulse response
KHz	kilo-hertz
LSI	linear shift invariant
LTI	linear time invariant
MA	moving average
MHz	mega-hertz
MSE	mean square error
PDF	probability density function
PMF	probability mass function
PSD	power spectral density
SNR	signal-to-noise ratio
WGN	white Gaussian noise
WSS	wide sense stationary

Appendix B

Assorted Math Facts and Formulas

An extensive summary of math facts and formulas can be found in [Gradshteyn and Ryzhik 1994].

B.1 Proof by Induction

To prove that a statement is true, for example,

$$\sum_{i=1}^{N} i = \frac{N}{2}(N+1) \tag{B.1}$$

by mathematical induction we proceed as follows:

1. Prove the statement is true for $N = 1$.

2. *Assume* the statement is true $N = n$ and prove that it therefore must be true for $N = n + 1$.

Obviously, (B.1) is true for $N = 1$ since $\sum_{i=1}^{1} i = 1$ and $(N/2)(N+1) = (1/2)(2) = 1$. Now assume it is true for $N = n$. Then for $N = n + 1$ we have

$$\begin{aligned}
\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + (n+1) \\
&= \frac{n}{2}(n+1) + (n+1) \qquad \text{(since it is true for } N = n\text{)} \\
&= \frac{n+1}{2}(n+2) \\
&= \frac{(n+1)}{2}[(n+1)+1]
\end{aligned}$$

which proves that it is also true for $N = n + 1$. By induction, since it is true for $N = n = 1$ from step 1, it must also be true for $N = (n + 1) = 2$ from step 2. And since it is true for $N = n = 2$, it must also be true for $N = n + 1 = 3$, etc.

B.2 Trigonometry

Some useful trigonometric identities are:

Fundamental

$$\sin^2 \alpha + \cos^2 \alpha = 1 \tag{B.2}$$

Sum of angles

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \tag{B.3}$$
$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \tag{B.4}$$

Double angle

$$\sin(2\alpha) = 2\sin\alpha\cos\alpha \tag{B.5}$$
$$\cos(2\alpha) = \cos^2 \alpha - \sin^2 \alpha = 2\cos^2 \alpha - 1 \tag{B.6}$$

Squared sine and cosine

$$\sin^2 \alpha = \frac{1}{2} - \frac{1}{2}\cos(2\alpha) \tag{B.7}$$
$$\cos^2 \alpha = \frac{1}{2} + \frac{1}{2}\cos(2\alpha) \tag{B.8}$$

Euler identities For $j = \sqrt{-1}$

$$\exp(j\alpha) = \cos\alpha + j\sin\alpha \tag{B.9}$$

$$\cos\alpha = \frac{\exp(j\alpha) + \exp(-j\alpha)}{2} \tag{B.10}$$
$$\sin\alpha = \frac{\exp(j\alpha) - \exp(-j\alpha)}{2j} \tag{B.11}$$

B.3 Limits

Alternative definition of exponential function

$$\lim_{M\to\infty} \left(1 + \frac{x}{M}\right)^M = \exp(x) \tag{B.12}$$

Taylor series expansion about the point $x = x_0$

$$g(x) = \sum_{i=0}^{\infty} \frac{g^{(i)}(x_0)}{i!}(x - x_0)^i \tag{B.13}$$

where $g^{(i)}(x_0)$ is the ith derivative of $g(x)$ evaluated at $x = x_0$ and $g^{(0)}(x_0) = g(x_0)$. As an example, consider $g(x) = \exp(x)$, which when expanded about $x = x_0 = 0$ yields

$$\exp(x) = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

B.4 Sums

Integers

$$\sum_{i=0}^{N-1} i = \frac{N(N-1)}{2} \tag{B.14}$$

$$\sum_{i=0}^{N-1} i^2 = \frac{N(N-1)(2N-1)}{6} \tag{B.15}$$

Real geometric series

$$\sum_{i=k}^{N-1} x^i = \frac{x^k(1 - x^{N-k})}{1 - x} \qquad (x \text{ is real}) \tag{B.16}$$

If $|x| < 1$, then

$$\sum_{i=k}^{\infty} x^i = \frac{x^k}{1 - x} \tag{B.17}$$

Complex geometric series

$$\sum_{i=k}^{N-1} z^i = \frac{z^k(1 - z^{N-k})}{1 - z} \qquad (z \text{ is complex}) \tag{B.18}$$

A special case is when $z = \exp(j\theta)$. Then

$$\begin{aligned} \sum_{i=0}^{N-1} \exp(j\theta) &= \frac{1 - \exp(jN\theta)}{1 - \exp(j\theta)} \\ &= \exp\left[j\left(\frac{N-1}{2}\right)\theta\right] \frac{\sin\left(\frac{N\theta}{2}\right)}{\sin\left(\frac{\theta}{2}\right)} \end{aligned} \tag{B.19}$$

If $|z| = |x + jy| = \sqrt{x^2 + y^2} < 1$, then as $N \to \infty$ (B.18) becomes

$$\sum_{i=k}^{\infty} z^i = \frac{z^k}{1-z} \tag{B.20}$$

Double sums

$$\sum_{i=1}^{M}\sum_{j=1}^{M}(x_i y_j) = \left(\sum_{i=1}^{M} x_i\right)\left(\sum_{j=1}^{M} y_j\right) \tag{B.21}$$

B.5 Calculus

Convergence of sum to integral

If $g(x)$ is a continuous function over $[a, b]$, then

$$\lim_{\Delta x \to 0} \sum_{i=0}^{M} g(x_i)\Delta x = \int_a^b g(x)dx \tag{B.22}$$

where $x_i = a + i\Delta x$ and $x_M = b$. Also, this shows how to approximate an integral by a sum.

Approximation of integral over small interval

$$\int_{x_0-\Delta x/2}^{x_0+\Delta x/2} g(x)dx \approx g(x_0)\Delta x \tag{B.23}$$

Differentiation of composite function

$$\left.\frac{dg(h(x))}{dx}\right|_{x=x_0} = \left.\frac{dg(u)}{du}\right|_{u=h(x_0)} \left.\frac{dh(x)}{dx}\right|_{x=x_0} \qquad \text{(chain rule)} \tag{B.24}$$

Change of integration variable

If $u = h(x)$, then

$$\int_a^b g(u)du = \int_{h^{-1}(a)}^{h^{-1}(b)} g(h(x))h'(x)dx \tag{B.25}$$

where $h'(x)$ is the derivative of $h(x)$ and $h^{-1}(\cdot)$ denotes the inverse function. This assumes that there is one solution to the equation $u = h(x)$ over the interval $a \leq u \leq b$.

Fundamental theorem of calculus

$$\frac{d}{dx}\int_{-\infty}^{x} g(t)dt = g(x) \tag{B.26}$$

Leibnitz's rule

$$\frac{d}{dy}\int_{h_1(y)}^{h_2(y)} g(x,y)dx = \int_{h_1(y)}^{h_2(y)} \frac{\partial}{\partial y} g(x,y)dx + g(h_2(y),y)\frac{dh_2(y)}{dy} - g(h_1(y),y)\frac{dh_1(y)}{dy} \tag{B.27}$$

Integration of even and odd functions An even function is defined as having the property $g(-x) = g(x)$, while an odd function has the property $g(-x) = -g(x)$. As a result,

$$\int_{-M}^{M} g(x)dx = 2\int_{0}^{M} g(x)dx \quad \text{for } g(x) \text{ an even function}$$

$$\int_{-M}^{M} g(x)dx = 0 \quad \text{for } g(x) \text{ an odd function}$$

Integration by parts

If U and V are both functions of x, then

$$\int UdV = UV - \int VdU \tag{B.28}$$

Dirac delta "function" or impulse

Denoted by $\delta(x)$ it is not really a function but a symbol that has the definition

$$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}$$

and

$$\int_a^b \delta(x)dx = \begin{cases} 1 & 0 \in [a^-, b^+] \\ 0 & \text{otherwise} \end{cases}$$

Some properties are for $u(x)$ the unit step function

$$\frac{du(x)}{dx} = \delta(x)$$

$$\int_{-\infty}^{x} \delta(t)dt = u(x)$$

Double integrals

$$\int_c^d \int_a^b g(x)h(y)dxdy = \left(\int_a^b g(x)dx\right)\left(\int_c^d h(y)dy\right) \tag{B.29}$$

References

Gradshteyn, I.S., I.M. Ryzhik, *Tables of Integrals, Series, and Products*, Fifth Ed., Academic Press, New York, 1994.

Appendix C

Linear and Matrix Algebra

Important results from linear and matrix algebra theory are reviewed in this appendix. It is assumed that the reader has had some exposure to matrices. For a more comprehensive treatment the books [Noble and Daniel 1977] and [Graybill 1969] are recommended.

C.1 Definitions

Consider an $M \times N$ matrix $\mathbf{A}$ with elements a_{ij}, $i = 1, 2, \ldots, M$; $j = 1, 2, \ldots, N$. A shorthand notation for describing $\mathbf{A}$ is

$$[\mathbf{A}]_{ij} = a_{ij}.$$

Likewise a shorthand notation for describing an $N \times 1$ vector $\mathbf{b}$ is

$$[\mathbf{b}]_i = b_i.$$

An $M \times N$ matrix $\mathbf{A}$ may multiply an $N \times 1$ vector $\mathbf{b}$ to yield a new $M \times 1$ vector $\mathbf{c}$ whose ith element is

$$c_i = \sum_{j=1}^{N} a_{ij} b_j \qquad i = 1, 2, \ldots, M.$$

Similarly, an $M \times N$ matrix $\mathbf{A}$ can multiply an $N \times L$ matrix $\mathbf{B}$ to yield an $M \times L$ matrix $\mathbf{C} = \mathbf{AB}$ whose (i, j) element is

$$c_{ij} = \sum_{k=1}^{N} a_{ik} b_{kj} \qquad i = 1, 2, \ldots, M; j = 1, 2, \ldots, L.$$

Vectors and matrices that can be multiplied together are said to be *conformable.*

The *transpose* of $\mathbf{A}$, which is denoted by $\mathbf{A}^T$, is defined as the $N \times M$ matrix with elements a_{ji} or

$$[\mathbf{A}^T]_{ij} = a_{ji}.$$

A *square* matrix is one for which $M = N$. A square matrix is *symmetric* if $\mathbf{A}^T = \mathbf{A}$ or $a_{ji} = a_{ij}$.

The *inverse* of a square $N \times N$ matrix is the square $N \times N$ matrix $\mathbf{A}^{-1}$ for which

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. If the inverse does not exist, then $\mathbf{A}$ is *singular*. Assuming the existence of the inverse of a matrix, the unique solution to a set of N simultaneous linear equations given in matrix form by $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is $N \times N$, $\mathbf{x}$ is $N \times 1$, and $\mathbf{b}$ is $N \times 1$, is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

The *determinant* of a square $N \times N$ matrix is denoted by $\det(\mathbf{A})$. It is computed as

$$\det(\mathbf{A}) = \sum_{j=1}^{N} a_{ij} C_{ij}$$

where

$$C_{ij} = (-1)^{i+j} D_{ij}.$$

D_{ij} is the determinant of the submatrix of $\mathbf{A}$ obtained by deleting the ith row and jth column and is termed the *minor* of a_{ij}. C_{ij} is the *cofactor* of a_{ij}. Note that any choice of i for $i = 1, 2, \ldots, N$ will yield the same value for $\det(\mathbf{A})$. A square $N \times N$ matrix is nonsingular if and only if $\det(\mathbf{A}) \neq 0$.

A *quadratic form* Q, which is a *scalar*, is defined as

$$Q = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} x_i x_j.$$

In defining the quadratic form it is assumed that $a_{ji} = a_{ij}$. This entails no loss in generality since any quadratic function may be expressed in this manner. Q may also be expressed as

$$Q = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

where $\mathbf{x} = [x_1\, x_2 \ldots x_N]^T$ and $\mathbf{A}$ is a square $N \times N$ matrix with $a_{ji} = a_{ij}$ or $\mathbf{A}$ is a symmetric matrix.

A square $N \times N$ matrix $\mathbf{A}$ is *positive semidefinite* if $\mathbf{A}$ is symmetric and

$$Q = \mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$$

for all $\mathbf{x}$. If the quadratic form is strictly positive for $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{A}$ is *positive definite*. When referring to a matrix as positive definite or positive semidefinite, it is always assumed that the matrix is symmetric.

A *partitioned* $M \times N$ matrix $\mathbf{A}$ is one that is expressed in terms of its submatrices. An example is the 2×2 partitioning

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}.$$

Each "element" $\mathbf{A}_{ij}$ is a submatrix of $\mathbf{A}$. The dimensions of the partitions are given as

$$\begin{bmatrix} K \times L & K \times (N-L) \\ (M-K) \times L & (M-K) \times (N-L) \end{bmatrix}.$$

C.2 Special Matrices

A *diagonal* matrix is a square $N \times N$ matrix with $a_{ij} = 0$ for $i \neq j$ or all elements not on the *principal diagonal* (the diagonal containing the elements a_{ii}) are zero. The elements a_{ij} for which $i \neq j$ are termed the *off-diagonal* elements. A diagonal matrix appears as

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{NN} \end{bmatrix}.$$

A diagonal matrix will sometimes be denoted by $\mathrm{diag}(a_{11}, a_{22}, \dots, a_{NN})$. The inverse of a diagonal matrix is found by simply inverting each element on the principal diagonal, assuming that $a_{ii} \neq 0$ for $i = 1, 2, \dots, N$ (which is necessary for invertibility).

A square $N \times N$ matrix is *orthogonal* if

$$\mathbf{A}^{-1} = \mathbf{A}^T.$$

For a matrix to be orthogonal the columns (and rows) must be orthonormal or if

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_N \end{bmatrix}$$

where $\mathbf{a}_i$ denotes the ith column, the conditions

$$\mathbf{a}_i^T \mathbf{a}_j = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$$

must be satisfied. Other "matrices" that can be constructed from vector operations on the $N \times 1$ vectors $\mathbf{x}$ and $\mathbf{y}$ are the *inner product*, which is defined as the scalar

$$\mathbf{x}^T \mathbf{y} = \sum_{i=1}^{N} x_i y_i$$

and the *outer product*, which is defined as the $N \times N$ matrix

$$\mathbf{x}\mathbf{y}^T = \begin{bmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_N \\ x_2y_1 & x_2y_2 & \cdots & x_2y_N \\ \vdots & \vdots & \ddots & \vdots \\ x_Ny_1 & x_Ny_2 & \cdots & x_Ny_N \end{bmatrix}.$$

C.3 Matrix Manipulation and Formulas

Some useful formulas for the algebraic manipulation of matrices are summarized in this section. For $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$ the following relationships are useful.

$$\begin{aligned} (\mathbf{A}^T)^{-1} &= (\mathbf{A}^{-1})^T \\ (\mathbf{AB})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} \\ \det(\mathbf{A}^T) &= \det(\mathbf{A}) \\ \det(c\mathbf{A}) &= c^N \det(\mathbf{A}) \quad (c \text{ a scalar}) \\ \det(\mathbf{AB}) &= \det(\mathbf{A})\det(\mathbf{B}) \\ \det(\mathbf{A}^{-1}) &= \frac{1}{\det(\mathbf{A})}. \end{aligned}$$

Also, for any conformable matrices (or vectors) we have

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T.$$

It is frequently necessary to determine the inverse of a matrix analytically. To do so one can make use of the following formula. The inverse of a square $N \times N$ matrix is

$$\mathbf{A}^{-1} = \frac{\mathbf{C}^T}{\det(\mathbf{A})}$$

where $\mathbf{C}$ is the square $N \times N$ matrix of cofactors of $\mathbf{A}$. The cofactor matrix is defined by

$$[\mathbf{C}]_{ij} = (-1)^{i+j} D_{ij}$$

where D_{ij} is the minor of a_{ij} obtained by deleting the ith row and jth column of $\mathbf{A}$.

Partitioned matrices may be manipulated according to the usual rules of matrix algebra by considering each submatrix as an element. For multiplication of partitioned matrices the submatrices that are multiplied together must be conformable. As an illustration, for 2×2 partitioned matrices

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{bmatrix}. \end{aligned}$$

Other useful relationships for partitioned matrices for an $M \times N$ matrix $\mathbf{A}$ and $N \times 1$ vectors $\mathbf{x}_i$ are

$$\begin{bmatrix} \mathbf{A}\mathbf{x}_1 & \mathbf{A}\mathbf{x}_2 & \ldots & \mathbf{A}\mathbf{x}_N \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_N \end{bmatrix} \tag{C.1}$$

which is a $M \times N$ matrix and

$$\begin{bmatrix} a_{11}\mathbf{x}_1 & a_{22}\mathbf{x}_2 & \ldots & a_{NN}\mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_N \end{bmatrix} \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{NN} \end{bmatrix} \tag{C.2}$$

which is an $N \times N$ matrix.

C.4 Some Properties of Positive Definite (Semidefinite) Matrices

Some useful properties of positive definite (semidefinite) matrices are:

1. A square $N \times N$ matrix $\mathbf{A}$ is positive definite if and only if the principal minors are all positive. (The ith principal minor is the determinant of the submatrix formed by deleting all rows and columns with an index greater than i.) If the principal minors are only nonnegative, then $\mathbf{A}$ is positive semidefinite.

2. If $\mathbf{A}$ is positive definite (positive semidefinite), then

  a. $\mathbf{A}$ is invertible (singular).

  b. the diagonal elements are positive (nonnegative).

  c. the determinant of $\mathbf{A}$, which is a principal minor, is positive (nonnegative).

C.5 Eigendecomposition of Matrices

An *eigenvector* of a square $N \times N$ matrix $\mathbf{A}$ is an $N \times 1$ vector $\mathbf{v}$ satisfying

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{C.3}$$

for some scalar λ, which may be complex. λ is the *eigenvalue* of $\mathbf{A}$ corresponding to the eigenvector $\mathbf{v}$. To determine the eigenvalues we must solve for the N λ's in $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, which is an Nth order polynomial in λ. Once the eigenvalues are found, the corresponding eigenvectors are determined from the equation $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$. It is assumed that the eigenvector is normalized to have unit length or $\mathbf{v}^T\mathbf{v} = 1$.

If $\mathbf{A}$ is symmetric, then one can always find N linearly independent eigenvectors, although they will not in general be unique. An example is the identity matrix for which any vector is an eigenvector with eigenvalue 1. If $\mathbf{A}$ is symmetric, then the eigenvectors corresponding to distinct eigenvalues are orthonormal or $\mathbf{v}_i^T\mathbf{v}_j = 0$ for $i \neq j$ and $\mathbf{v}_i^T\mathbf{v}_j = 1$ for $i = j$, and the eigenvalues are real. If, furthermore, the matrix is positive definite (positive semidefinite), then the eigenvalues are positive (nonnegative).

The defining relation of (C.3) can also be written as (using (C.1) and (C.2))

$$\begin{bmatrix} \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \ldots & \mathbf{A}\mathbf{v}_N \end{bmatrix} = \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \ldots & \lambda_N\mathbf{v}_n \end{bmatrix}$$

or

$$\mathbf{AV} = \mathbf{V\Lambda} \tag{C.4}$$

where

$$\begin{aligned} \mathbf{V} &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{bmatrix} \\ \mathbf{\Lambda} &= \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n). \end{aligned}$$

If $\mathbf{A}$ is symmetric so that the eigenvectors corresponding to distinct eigenvalues are orthonormal and the remaining eigenvectors are chosen to yield an orthonormal eigenvector set, then $\mathbf{V}$ is an orthogonal matrix. As such, its inverse is $\mathbf{V}^T$, so that (C.4) becomes

$$\mathbf{A} = \mathbf{V\Lambda V}^T$$

Also, the inverse is easily determined as

$$\begin{aligned} \mathbf{A}^{-1} &= \mathbf{V}^{T^{-1}}\mathbf{\Lambda}^{-1}\mathbf{V}^{-1} \\ &= \mathbf{V\Lambda}^{-1}\mathbf{V}^T. \end{aligned}$$

References

Graybill, F.A., *Introduction to Matrices with Applications in Statistics*, Wadsworth, Belmont, CA, 1969.

Noble, B., Daniel, J.W., *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

Appendix D

Summary of Signals, Linear Transforms, and Linear Systems

In this appendix we summarize the important concepts and formulas for discrete-time signal and system analysis. This material is used in Chapters 18–20. Some examples are given so that the reader unfamiliar with this material should try to verify the example results. For a more comprehensive treatment the books [Jackson 1991], [Oppenheim, Willsky, and Nawab 1997], [Poularikis and Seeley 1985] are recommended.

D.1 Discrete-Time Signals

A discrete-time signal is a *sequence* $x[n]$ for $n = \ldots, -1, 0, 1, \ldots$. It is defined only for the integers. Some important signals are:

a. *Unit impulse* – $x[n] = 1$ for $n = 0$ and $x[n] = 0$ for $n \neq 0$. It is also denoted by $\delta[n]$.

b. *Unit step* – $x[n] = 1$ for $n \geq 0$ and $x[n] = 0$ for $n < 0$. It is also denoted by $u[n]$.

c. *Real sinusoid* – $x[n] = A\cos(2\pi f_0 n + \theta)$ for $-\infty < n < \infty$, where A is the amplitude (must be nonnegative), f_0 is the frequency in cycles per sample and must be in the interval $0 < f_0 < 1/2$, and θ is the phase in radians.

d. *Complex sinusoid* – $x[n] = A\exp(j2\pi f_0 n + \theta)$ for $-\infty < n < \infty$, where A is the amplitude (must be nonnegative), f_0 is the frequency in cycles per sample and must be in the interval $-1/2 < f_0 < 1/2$, and θ is the phase in radians.

e. *Exponential* – $x[n] = a^n u[n]$

Note that any sequence can be written as a linear combination of unit impulses that are weighted by $x[k]$ and shifted in time as $\delta[n-k]$ to form

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]. \tag{D.1}$$

For example, $a^n u[n] = \delta[n] + a\delta[n-1] + a^2\delta[n-2] + \cdots$.

Some special signals are defined next.

a. A signal is *causal* if $x[n] = 0$ for $n < 0$, for example, $x[n] = u[n]$.

b. A signal is *anticausal* if $x[n] = 0$ for $n > 0$, for example, $x[n] = u[-n]$.

c. A signal is *even* if $x[-n] = x[n]$ or it is symmetric about $n = 0$, for example, $x[n] = \cos(2\pi f_0 n)$.

d. A signal is *odd* if $x[-n] = -x[n]$ or it is antisymmetric about $n = 0$, for example, $x[n] = \sin(2\pi f_0 n)$.

e. A signal is *stable* if $\sum_{n=-\infty}^{\infty} |x[n]| < \infty$ (also called *absolutely summable*), for example, $x[n] = (1/2)^n u[n]$.

D.2 Linear Transforms

D.2.1 Discrete-Time Fourier Transforms

The *discrete-time Fourier transform* $X(f)$ of a discrete-time signal $x[n]$ is defined as

$$X(f) = \sum_{n=-\infty}^{\infty} x[n]\exp(-j2\pi f n) \qquad -1/2 \le f \le 1/2. \tag{D.2}$$

An example is $x[n] = (1/2)^n u[n]$ for which $X(f) = 1/(1-(1/2)\exp(-j2\pi f))$. It converts a discrete-time signal into a complex function of f, where f is called the *frequency* and is measured in cycles per sample. The operation of taking the Fourier transform of a signal is denoted by $\mathcal{F}\{x[n]\}$ and the signal and its Fourier transform are referred to as a *Fourier transform pair.* The latter relationship is usually denoted by $x[n] \Leftrightarrow X(f)$. The discrete-time Fourier transform is *periodic in frequency with period one* and for this reason we need only consider the frequency interval $[-1/2, 1/2]$. Since the Fourier transform is a complex function of frequency, it can be represented by the two real functions

$$\begin{aligned}
|X(f)| &= \sqrt{\left(\sum_{n=-\infty}^{\infty} x[n]\cos(2\pi f n)\right)^2 + \left(\sum_{n=-\infty}^{\infty} x[n]\sin(2\pi f n)\right)^2} \\
\phi(f) &= \arctan\frac{-\sum_{n=-\infty}^{\infty} x[n]\sin(2\pi f n)}{\sum_{n=-\infty}^{\infty} x[n]\cos(2\pi f n)}
\end{aligned}$$

Signal name	$x[n]$	$X(f)\ (-\frac{1}{2} \le f \le \frac{1}{2})$
Unit impulse	$\delta[n] = \begin{cases} 1 & n=0 \\ 0 & n \neq 0 \end{cases}$	1
Real sinusoid	$\cos(2\pi f_0 n)$	$\frac{1}{2}\delta(f+f_0) + \frac{1}{2}\delta(f-f_0)$
Complex sinusoid	$\exp(j2\pi f_0 n)$	$\delta(f-f_0)$
Exponential	$a^n u[n]$	$\frac{1}{1-a\exp(-j2\pi f)} \quad \lvert a\rvert < 1$
Double-sided exponential	$a^{\lvert n\rvert}$	$\frac{1-a^2}{1+a^2-2a\cos(2\pi f)} \quad \lvert a\rvert < 1$

Table D.1: Discrete-time Fourier transform pairs.

which are called the *magnitude* and *phase*, respectively. For example, if $x[n] = (1/2)^n u[n]$, then

$$\begin{aligned} |X(f)| &= \frac{1}{\sqrt{5/4 - \cos(2\pi f)}} \\ \phi(f) &= -\arctan \frac{\frac{1}{2}\sin(2\pi f)}{1 - \frac{1}{2}\cos(2\pi f)}. \end{aligned}$$

Note that the magnitude is an even function or $|X(-f)| = |X(f)|$ and the phase is an odd function or $\phi(-f) = -\phi(f)$. Some Fourier transform pairs are given in Table D.1. Some important properties of the discrete-time Fourier transform are:

a. Linearity – $\mathcal{F}\{ax[n] + by[n]\} = aX(f) + bY(f)$

b. Time shift – $\mathcal{F}\{x[n-n_0]\} = \exp(-j2\pi f n_0)X(f)$

c. Modulation – $\mathcal{F}\{\cos(2\pi f_0 n)x[n]\} = \frac{1}{2}X(f+f_0) + \frac{1}{2}X(f-f_0)$

d. Time reversal – $\mathcal{F}\{x[-n]\} = X^*(f)$

e. Symmetry – if $x[n]$ is even, then $X(f)$ is even and real, and if $x[n]$ is odd, then $X(f)$ is odd and purely imaginary.

f. Energy – the energy defined as $\sum_{n=-\infty}^{\infty} x^2[n]$ can be found from the Fourier transform using *Parseval's theorem*

$$\sum_{n=-\infty}^{\infty} x^2[n] = \int_{-\frac{1}{2}}^{\frac{1}{2}} |X(f)|^2 df.$$

g. Inner product – as an extension of Parseval's theorem we have

$$\sum_{n=-\infty}^{\infty} x[n]y[n] = \int_{-\frac{1}{2}}^{\frac{1}{2}} X^*(f)Y(f)df.$$

Two signals $x[n]$ and $y[n]$ are said to be *convolved* together to yield a new signal $z[n]$ if

$$z[n] = \sum_{k=-\infty}^{\infty} x[k]y[n-k] \qquad -\infty < n < \infty.$$

As an example, if $x[n] = u[n]$ and $y[n] = u[n]$, then $z[n] = (n+1)u[n]$. The operation of convolving two signals together is called *convolution* and is implemented using a *convolution sum.* It is denoted by $x[n] \star y[n]$. The operation is commutative in that $x[n] \star y[n] = y[n] \star x[n]$ so that an equivalent form is

$$z[n] = \sum_{k=-\infty}^{\infty} y[k]x[n-k] \qquad -\infty < n < \infty.$$

As an example, if $y[n] = \delta[n-n_0]$, then it is easily shown that $x[n] \star \delta[n-n_0] = \delta[n-n_0] \star x[n] = x[n-n_0]$. The most important property of convolution is that two signals that are convolved together produce a signal whose Fourier transform is the product of the signals' Fourier transforms or

$$\mathcal{F}\{x[n] \star y[n]\} = X(f)Y(f).$$

Two signals $x[n]$ and $y[n]$ are said to be *correlated together* to yield a new signal $z[n]$ if

$$z[n] = \sum_{k=-\infty}^{\infty} x[k]y[k+n] \qquad -\infty < n < \infty.$$

The Fourier transform of $z[n]$ is $X^*(f)Y(f)$. The sequence $z[n]$ is also called the *deterministic cross-correlation.* If $x[n] = y[n]$, then $z[n]$ is called the *deterministic autocorrelation* and its Fourier transform is $|X(f)|^2$.

The discrete-time signal may be recovered from its Fourier transform by using the *discrete-time inverse Fourier transform*

$$x[n] = \int_{-\frac{1}{2}}^{\frac{1}{2}} X(f)\exp(j2\pi fn)df \qquad -\infty < n < \infty. \tag{D.3}$$

As an example, if $X(f) = \frac{1}{2}\delta(f+f_0) + \frac{1}{2}\delta(f-f_0)$, then the integral yields $x[n] = \cos(2\pi f_0 n)$. It also has the interpretation that a discrete-time signal $x[n]$ may be thought of as a sum of complex sinusoids $X(f)\exp(j2\pi fn)\Delta f$ for $-1/2 \leq f \leq 1/2$ with amplitude $|X(f)|\Delta f$ and phase $\angle X(f)$. There is a separate sinusoid for each frequency f, and the total number of sinusoids is uncountable.

D.2.2 Numerical Evaluation of Discrete-Time Fourier Transforms

The discrete-time Fourier transform of a signal $x[n]$, which is nonzero only for $n = 0, 1, \ldots, N-1$, is given by

$$X(f) = \sum_{n=0}^{N-1} x[n] \exp(-j2\pi fn) \qquad -1/2 \leq f \leq 1/2. \tag{D.4}$$

Such a signal is said to be *time-limited*. Since the Fourier transform is periodic with period one, we can equivalently evaluate it over the interval $0 \leq f \leq 1$. Then, if we desire the Fourier transform for $-1/2 \leq f' < 0$, we use the previously evaluated $X(f)$ with $f = f' + 1$. To numerically evaluate the Fourier transform we therefore can use the frequency interval $[0, 1]$ and compute samples of $X(f)$ for $f = 0, 1/N, 2/N, \ldots, (N-1)/N$. This yields the *discrete Fourier transform* (DFT) which is defined as

$$X[k] = X(f)|_{f=k/N} = \sum_{n=0}^{N-1} x[n] \exp\left(-j2\pi(k/N)n\right) \qquad k = 0, 1, \ldots, N-1.$$

Since there are only N time samples, we may wish to compute more *frequency samples* since $X(f)$ is a *continuous* function of frequency. To do so we can *zero pad* the time samples with zeros to yield a new signal $x'[n]$ of length $M > N$ with samples $\{x[0], x[1], \ldots, x[N-1], 0, 0, \ldots, 0\}$. This new signal $x'[n]$ will consist of N time samples and $M - N$ zeros so that the DFT will compute more finely spaced frequency samples as

$$\begin{aligned} X[k] &= X(f)|_{f=k/M} = \sum_{n=0}^{M-1} x'[n] \exp\left(-j2\pi(k/M)n\right) \qquad k = 0, 1, \ldots, M-1 \\ &= \sum_{n=0}^{N-1} x[n] \exp\left(-j2\pi(k/M)n\right) \qquad k = 0, 1, \ldots, M-1. \end{aligned}$$

The actual DFT is computed using the fast Fourier transform (FFT), which is an algorithm used to reduce the computation.

The inverse Fourier transform of an infinite length causal sequence can be *approximated* using an *inverse DFT* as

$$\begin{aligned} x[n] &= \int_{-\frac{1}{2}}^{\frac{1}{2}} X(f) \exp(j2\pi fn) df = \int_0^1 X(f) \exp(j2\pi fn) df \\ &\approx \frac{1}{M} \sum_{k=0}^{M-1} X[k] \exp\left(j2\pi(k/M)n\right) \qquad n = 0, 1, \ldots, M-1. \end{aligned} \tag{D.5}$$

One should choose M large. The actual inverse DFT is computed using the inverse FFT.

D.2.3 z-Transforms

The z-transform of a discrete-time signal $x[n]$ is defined as

$$\mathcal{X}(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n} \tag{D.6}$$

where z is a complex variable that takes on values for which $|\mathcal{X}(z)| < \infty$. As an example, if $x[n] = (1/2)^n u[n]$, then

$$\mathcal{X}(z) = \frac{1}{1 - \frac{1}{2}z^{-1}} \qquad |z| > \frac{1}{2}. \tag{D.7}$$

The operation of taking the z-transform is indicated by $\mathcal{Z}\{x[n]\}$. Some important properties of the z-transform are:

a. Linearity – $\mathcal{Z}\{ax[n] + by[n]\} = a\mathcal{X}(z) + b\mathcal{Y}(z)$

b. Time shift – $\mathcal{Z}\{x[n - n_0]\} = z^{-n_0}\mathcal{X}(z)$

c. Convolution – $\mathcal{Z}\{x[n] \star y[n]\} = \mathcal{X}(z)\mathcal{Y}(z)$.

Assuming that the z-transform converges on the unit circle, the discrete-time Fourier transform is given by

$$X(f) = \mathcal{X}(z)|_{z=\exp(j2\pi f)} \tag{D.8}$$

as is seen by comparing (D.6) to (D.2). As an example, if $x[n] = (1/2)^n u[n]$, then from (D.7)

$$X(f) = \frac{1}{1 - \frac{1}{2}\exp(-j2\pi f)}$$

since $\mathcal{X}(z)$ converges for $|z| = |\exp(j2\pi f)| = 1 > 1/2$.

D.3 Discrete-Time Linear Systems

A discrete-time system takes an input signal $x[n]$ and produces an output signal $y[n]$. The transformation is symbolically represented as $y[n] = \mathcal{L}\{x[n]\}$. The system is *linear* if $\mathcal{L}\{ax[n] + by[n]\} = a\mathcal{L}\{x[n]\} + b\mathcal{L}\{y[n]\}$. A system is defined to be *shift invariant* if $\mathcal{L}\{x[n - n_0]\} = y[n - n_0]$. If the system is *linear and shift invariant* (LSI), then the output is easily found if we know the output to a unit impulse. To see this we compute the output of the system as

$$\begin{aligned}
y[n] &= \mathcal{L}\{x[n]\} \\
&= \mathcal{L}\left\{\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\right\} && \text{(using (D.1))} \\
&= \sum_{k=-\infty}^{\infty} x[k]\mathcal{L}\{\delta[n-k]\} && \text{(linearity)} \\
&= \sum_{k=-\infty}^{\infty} x[k]\ \mathcal{L}\{\delta[n]\}|_{n\to n-k} && \text{(shift invariance)} \\
&= \sum_{k=-\infty}^{\infty} x[k]h[n-k]
\end{aligned}$$

where $h[n] = \mathcal{L}\{\delta[n]\}$ is called the *impulse response* of the system. Note that $y[n] = x[n] \star h[n] = h[n] \star x[n]$ and so the output of the LSI system is also given by the convolution sum

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k]. \tag{D.9}$$

A *causal* system is defined as one for which $h[k] = 0$ for $k < 0$ since then the output depends only on the present input $x[n]$ and the past inputs $x[n-k]$ for $k \geq 1$. The system is said to be *stable* if

$$\sum_{k=-\infty}^{\infty} |h[k]| < \infty.$$

If this condition is satisfied, then a bounded input signal or $|x[n]| < \infty$ for $-\infty < n < \infty$ will always produce a bounded output signal or $|y[n]| < \infty$ for $-\infty < n < \infty$. As an example, the LSI system with impulse response $h[k] = (1/2)^k u[k]$ is stable but not the one with impulse response $h[k] = u[k]$. The latter system will produce the unbounded output $y[n] = (n+1)u[n]$ for the bounded input $x[n] = u[n]$ since $u[n] \star u[n] = (n+1)u[n]$.

Since for an LSI system $y[n] = h[n] \star x[n]$, it follows from the properties of z-transforms that $\mathcal{Y}(z) = \mathcal{H}(z)\mathcal{X}(z)$, where $\mathcal{H}(z)$ is the z-transform of the impulse response. As a result, we have that

$$\mathcal{H}(z) = \frac{\mathcal{Y}(z)}{\mathcal{X}(z)} = \frac{\text{Output } z\text{-transform}}{\text{Input } z\text{-transform}} \tag{D.10}$$

and $\mathcal{H}(z)$ is called the *system function*. Note that since it is the z-transform of the impulse response $h[n]$ we have

$$\mathcal{H}(z) = \sum_{k=-\infty}^{\infty} h[n]z^{-n}. \tag{D.11}$$

If the input to an LSI system is a complex sinusoid, $x[n] = \exp(j2\pi f_0 n)$, then the output is from (D.9)

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k] \exp[j2\pi f_0(n-k)] \\ &= \underbrace{\sum_{k=-\infty}^{\infty} h[k] \exp(-j2\pi f_0 k)}_{H(f_0)} \exp(j2\pi f_0 n). \end{aligned} \tag{D.12}$$

It is seen that the output is also a complex sinusoid with the same frequency but multiplied by the Fourier transform of the impulse response evaluated at the sinusoidal frequency. Hence, $H(f)$ is called the *frequency response.* Also, from (D.12) the frequency response is obtained from the system function (see (D.11)) by letting $z = \exp(j2\pi f)$. Finally, note that the frequency response is the discrete-time Fourier transform of the impulse response. As an example, if $h[n] = (1/2)^n u[n]$, then

$$\mathcal{H}(z) = \frac{1}{1 - \frac{1}{2}z^{-1}}$$

and

$$H(f) = \mathcal{H}(\exp(j2\pi f)) = \frac{1}{1 - \frac{1}{2}\exp(-j2\pi f)}.$$

The *magnitude response* of the LSI system is defined as $|H(f)|$ and the *phase response* as $\angle H(f)$.

As we have seen, LSI systems can be characterized by the equivalent descriptions: impulse response, system function, or frequency response. This means that *given one of these descriptions the output can be determined for any input.* LSI systems can also be characterized by linear difference equations with constant coefficients. Some examples are

$$\begin{aligned} y_1[n] &= x[n] - bx[n-1] \\ y_2[n] &= ay_2[n-1] + x[n] \\ y_3[n] &= ay_3[n-1] + x[n] - bx[n-1] \end{aligned}$$

and more generally

$$y[n] = \sum_{k=1}^{p} a[k]y[n-k] + x[n] - \sum_{k=1}^{q} b[k]x[n-k]. \tag{D.13}$$

The system function is found by taking the z-transform of both sides of the difference equations and using (D.10) to yield

$$\begin{aligned}\mathcal{Y}_1(z) &= \mathcal{X}(z) - bz^{-1}\mathcal{X}(z) \Rightarrow \mathcal{H}_1(z) = 1 - bz^{-1} \\ \mathcal{Y}_2(z) &= az^{-1}\mathcal{Y}_2(z) + \mathcal{X}(z) \Rightarrow \mathcal{H}_2(z) = \frac{1}{1 - az^{-1}} \\ \mathcal{Y}_3(z) &= az^{-1}\mathcal{Y}_3(z) + \mathcal{X}(z) - bz^{-1}\mathcal{X}(z) \Rightarrow \mathcal{H}_3(z) = \frac{1 - bz^{-1}}{1 - az^{-1}}\end{aligned}$$

and the frequency response is obtained using $H(f) = \mathcal{H}(\exp(j2\pi f))$. More generally, for the LSI system whose difference equation description is given by (D.13) we have

$$\mathcal{H}(z) = \frac{1 - \sum_{k=1}^{q} b[k]z^{-k}}{1 - \sum_{k=1}^{p} a[k]z^{-k}}. \tag{D.14}$$

The impulse response is obtained by taking the inverse z-transform of the system function to yield for the previous examples

$$\begin{aligned}h_1[n] &= \begin{cases} 1 & n = 0 \\ -b & n = 1 \\ 0 & \text{otherwise} \end{cases} \\ h_2[n] &= a^n u[n] && \text{(assuming system is causal)} \\ h_3[n] &= a^n u[n] - ba^{n-1}u[n-1] && \text{(assuming system is causal).}\end{aligned}$$

The impulse response could also be obtained by letting $x[n] = \delta[n]$ in the difference equations and setting $y[-1] = 0$, due to causality, and *recursing* the difference equation. For example, if the difference equation is $y[n] = (1/2)y[n-1] + x[n]$, then by definition the impulse response satisfies the equation $h[n] = (1/2)h[n-1] + \delta[n]$. By recursing this we obtain

$$\begin{aligned}h[0] &= \tfrac{1}{2}h[-1] + \delta[0] = 1 && \text{(since } h[-1] = 0 \text{ due to causality)} \\ h[1] &= \tfrac{1}{2}h[0] + \delta[1] = \tfrac{1}{2} && \text{(since } \delta[n] = 0 \text{ for } n \geq 1) \\ h[2] &= \tfrac{1}{2}h[1] = \tfrac{1}{4} \\ & \text{etc.}\end{aligned}$$

and so in general we have the impulse response $h[n] = (1/2)^n u[n]$. The system with impulse response $h_1[n]$ is called a *finite impulse response* (FIR) system while those of $h_2[n]$ and $h_3[n]$ are called *infinite impulse response* (IIR) systems. The terminology refers to the number of nonzero samples of the impulse response.

For the system function $H_3(z) = (1 - bz^{-1})/(1 - az^{-1})$, the value of z for which the numerator is zero is called a *zero* and the value of z for which the denominator is zero is called a *pole*. In this case the system function has one zero at $z = b$ and one pole at $z = a$. For the system to be stable, assuming it is causal, *all the poles of the system function must be within the unit circle of the z-plane.* Hence, for stability we require $|a| < 1$. The zeros may lie anywhere in the z-plane. For a *second-order* system function (let $p = 2$ and $q = 0$ in (D.14)) given as

$$\mathcal{H}(z) = \frac{1}{1 - a[1]z^{-1} - a[2]z^{-2}}$$

the poles, assuming they are complex, are located at $z = r\exp(\pm j\theta)$. Hence, for stability we require $r < 1$ and we note that since the poles are the z values for which the denominator polynomial is zero, we have

$$1 - a[1]z^{-1} - a[2]z^{-2} = z^{-2}(z - r\exp(j\theta))(z - r\exp(-j\theta)).$$

Therefore, the coefficients are related to the complex poles as

$$\begin{aligned} a[1] &= 2r\cos(\theta) \\ a[2] &= -r^2 \end{aligned}$$

which puts restrictions on the possible values of $a[1]$ and $a[2]$. As an example, the coefficients $a[1] = 0$, $a[2] = -1/4$ produce a stable filter but not $a[1] = 0$, $a[2] = -2$.

An LSI system whose frequency response is

$$H(f) = \begin{cases} 1 & |f| \leq B \\ 0 & |f| > B \end{cases}$$

is said to be an *ideal lowpass filter.* It passes complex sinusoids undistorted if their frequency is $|f| \leq B$ but nullifies ones with a higher frquency. The band of positive frequencies from $f = 0$ to $f = B$ is called the *passband* and the band of positive frequencies for which $f > B$ is called the *stopband.*

D.4 Continuous-Time Signals

A continuous-time signal is a *function of time* $x(t)$ for $-\infty < t < \infty$. Some important signals are:

a. *Unit impulse* – It is denoted by $\delta(t)$. An impulse $\delta(t)$, also called the *Dirac delta function*, is defined as the limit of a very narrow pulse as the pulsewidth goes to zero and the pulse amplitude goes to infinity, such that the overall area remains at one. Therefore, if we define a very narrow pulse as

$$x_T(t) = \begin{cases} \frac{1}{T} & |t| \leq T/2 \\ 0 & |t| > T/2 \end{cases}$$

then the unit impulse is defined as

$$\delta(t) = \lim_{T \to 0} x_T(t).$$

The impulse has the important *sifting property* that if $x(t)$ is continuous at $t = t_0$, then

$$\int_{-\infty}^{\infty} x(t)\delta(t - t_0)dt = x(t_0).$$

b. *Unit step* – $x(t) = 1$ for $t \geq 0$ and $x(t) = 0$ for $t < 0$. It is also denoted by $u(t)$.

c. *Real sinusoid* – $x(t) = A\cos(2\pi F_0 t + \theta)$ for $-\infty < t < \infty$, where A is the amplitude (must be nonnegative), F_0 is the frequency in Hz (cycles per second), and θ is the phase in radians.

d. *Complex sinusoid* – $x(t) = A\exp(j2\pi F_0 t + \theta)$ for $-\infty < t < \infty$, with the amplitude, frequency, and phase taking on same values as for real sinusoid.

e. *Exponential* – $x(t) = \exp(at)u(t)$

f. *Pulse* – $x(t) = 1$ for $|t| \leq T/2$ and $x(t) = 0$ for $|t| > T/2$.

Some special signals are defined next.

a. A signal is *causal* if $x(t) = 0$ for $t < 0$, for example, $x(t) = u(t)$.

b. A signal is *anticausal* if $x(t) = 0$ for $t > 0$, for example, $x(t) = u(-t)$.

c. A signal is *even* if $x(-t) = x(t)$ or it is symmetric about $t = 0$, for example, $x(t) = \cos(2\pi F_0 t)$.

d. A signal is *odd* if $x(-t) = -x(t)$ or it is antisymmetric about $t = 0$, for example, $x(t) = \sin(2\pi F_0 t)$.

e. A signal is *stable* if $\int_{-\infty}^{\infty} |x(t)|dt < \infty$ (also called *absolutely integrable*), for example, $x(t) = \exp(-t)u(t)$.

D.5 Linear Transforms

D.5.1 Continuous-Time Fourier Transforms

The *continuous-time Fourier transform* $X(F)$ of a continuous-time signal $x(t)$ is defined as

$$X(F) = \int_{-\infty}^{\infty} x(t)\exp(-j2\pi Ft)dt \qquad -\infty < F < \infty. \tag{D.15}$$

An example is $x(t) = \exp(-t)u(t)$ for which $X(F) = 1/(1 + j2\pi F)$. It converts a continuous-time signal into a complex function of F, where F is called the *frequency* and is measured in Hz (cycles per second). The operation of taking the Fourier transform of a signal is denoted by $\mathcal{F}\{x(t)\}$ and the signal and its Fourier transform are referred to as a *Fourier transform pair.* The latter relationship is usually

Signal name	$x(t)$	$X(F)$
Unit impulse	$\delta(t) = \begin{cases} \infty & t=0 \\ 0 & t \neq 0 \end{cases}$	1
Real sinusoid	$\cos(2\pi F_0 t)$	$\frac{1}{2}\delta(F+F_0) + \frac{1}{2}\delta(F-F_0)$
Complex sinusoid	$\exp(j2\pi F_0 t)$	$\delta(F-F_0)$
Exponential	$\exp(-at)u(t)$	$\frac{1}{a+j2\pi F}$ $\quad a>0$
Pulse	$= \begin{cases} 1 & \|t\| \leq T/2 \\ 0 & \|t\| > T/2 \end{cases}$	$T\frac{\sin(\pi FT)}{\pi FT}$

Table D.2: Continuous-time Fourier transform pairs.

denoted by $x(t) \Leftrightarrow X(F)$. Note that the magnitude of $X(F)$ is an even function or $|X(-F)| = |X(F)|$ and the phase is an odd function or $\phi(-F) = -\phi(F)$. Some Fourier transform pairs are given in Table D.2.

Some important properties of the continuous-time Fourier transform are:

a. Linearity – $\mathcal{F}\{ax(t)+by(t)\} = aX(F)+bY(F)$

b. Time shift – $\mathcal{F}\{x(t-t_0)\} = \exp(-j2\pi F t_0)X(F)$

c. Modulation – $\mathcal{F}\{\cos(2\pi F_0 t)x(t)\} = \frac{1}{2}X(F+F_0) + \frac{1}{2}X(F-F_0)$

d. Time reversal – $\mathcal{F}\{x(-t)\} = X^*(F)$

e. Symmetry – if $x(t)$ is even, then $X(F)$ is even and real, and if $x(t)$ is odd, then $X(F)$ is odd and purely imaginary.

f. Energy – the energy defined as $\int_{-\infty}^{\infty} x^2(t)dt$ can be found from the Fourier transform using *Parseval's theorem*

$$\int_{-\infty}^{\infty} x^2(t)dt = \int_{-\infty}^{\infty} |X(F)|^2 dF.$$

g. Inner product – as an extension of Parseval's theorem we have

$$\int_{-\infty}^{\infty} x(t)y(t)dt = \int_{-\infty}^{\infty} X^*(F)Y(F)dF.$$

Two signals $x(t)$ and $y(t)$ are said to be *convolved* together to yield a new signal $z(t)$ if

$$z(t) = \int_{-\infty}^{\infty} x(\tau)y(t-\tau)d\tau \qquad -\infty < t < \infty.$$

As an example, if $x(t) = u(t)$ and $y(t) = u(t)$, then $z(t) = tu(t)$. The operation of convolving two signals together is called *convolution* and is implemented using a *convolution integral.* It is denoted by $x(t) \star y(t)$. The operation is commutative in that $x(t) \star y(t) = y(t) \star x(t)$ so that an equivalent form is

$$z(t) = \int_{-\infty}^{\infty} y(\tau)x(t-\tau)d\tau \qquad -\infty < t < \infty.$$

As an example, if $y(t) = \delta(t-t_0)$, then it is easily shown that $x(t) \star \delta(t-t_0) = \delta(t-t_0) \star x(t) = x(t-t_0)$. The most important property of convolution is that two signals that are convolved together produce a signal whose Fourier transform is the product of the signals' Fourier transforms or

$$\mathcal{F}\{x(t) \star y(t)\} = X(F)Y(F).$$

The continuous-time signal may be recovered from its Fourier transform by using the *continuous-time inverse Fourier transform*

$$x(t) = \int_{-\infty}^{\infty} X(F)\exp(j2\pi Ft)dF \qquad -\infty < t < \infty. \tag{D.16}$$

As an example, if $X(F) = \frac{1}{2}\delta(F+F_0) + \frac{1}{2}\delta(F-F_0)$, then the integral yields $x(t) = \cos(2\pi F_0 t)$. It also has the interpretation that a continuous-time signal $x(t)$ may be thought of as a sum of complex sinusoids $X(F)\exp(j2\pi Ft)\Delta F$ for $-\infty < F < \infty$ with amplitude $|X(F)|\Delta F$ and phase $\angle X(F)$. There is a separate sinusoid for each frequency F, and the total number of sinusoids is uncountable.

D.6 Continuous-Time Linear Systems

A continuous-time system takes an input signal $x(t)$ and produces an output signal $y(t)$. The transformation is symbolically represented as $y(t) = \mathcal{L}\{x(t)\}$. The system is *linear* if $\mathcal{L}\{ax(t) + by(t)\} = a\mathcal{L}\{x(t)\} + b\mathcal{L}\{y(t)\}$. A system is defined to be *time invariant* if $\mathcal{L}\{x(t-t_0)\} = y(t-t_0)$. If the system is *linear and time invariant* (LTI), then the output is easily found if we know the output to a unit impulse. It is given by the convolution integral

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau \tag{D.17}$$

where $h(t) = \mathcal{L}\{\delta(t)\}$ is called the *impulse response* of the system. A *causal* system is defined as one for which $h(\tau) = 0$ for $\tau < 0$ since then the output depends only on the present input $x(t)$ and the past inputs $x(t-\tau)$ for $\tau > 0$. The system is said to be *stable* if

$$\int_{-\infty}^{\infty} |h(\tau)|d\tau < \infty.$$

If this condition is satisfied, then a bounded input signal or $|x(t)| < \infty$ for $-\infty < t < \infty$ will always produce a bounded output signal or $|y(t)| < \infty$ for $-\infty < t < \infty$. As an example, the LTI system with impulse response $h(\tau) = \exp(-\tau)u(\tau)$ is stable but not the one with impulse response $h(\tau) = u(\tau)$. The latter system will produce the unbounded output $y(t) = tu(t)$ for the bounded input $x(t) = u(t)$ since $u(t) \star u(t) = tu(t)$.

If the input to an LTI system is a complex sinusoid, $x(t) = \exp(j2\pi F_0 t)$, then the output is from (D.17)

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)\exp[j2\pi F_0(t-\tau)]d\tau \\ &= \underbrace{\int_{-\infty}^{\infty} h(\tau)\exp(-j2\pi F_0\tau)d\tau}_{H(F_0)} \exp(j2\pi F_0 t). \end{aligned} \tag{D.18}$$

It is seen that the output is also a complex sinusoid with the same frequency but multiplied by the Fourier transform of the impulse response evaluated at the sinusoidal frequency. Hence, $H(F)$ is called the *frequency response.* Finally, note that the frequency response is the continuous-time Fourier transform of the impulse response. As an example, if $h(t) = \exp(-at)u(t)$, then for $a > 0$

$$H(F) = \frac{1}{a + j2\pi F}.$$

The *magnitude response* of the LSI system is defined as $|H(F)|$ and the *phase response* as $\angle H(F)$.

An LTI system whose frequency response is

$$H(F) = \begin{cases} 1 & |F| \le W \\ 0 & |F| > W \end{cases}$$

is said to be an *ideal lowpass filter.* It passes complex sinusoids undistorted if their frequency is $|F| \le W$ Hz but nullifies ones with a higher frquency. The band of positive frequencies from $F = 0$ to $F = W$ is called the *passband* and the band of positive frequencies for which $F > W$ is called the *stopband.*

References

Jackson, L.B., *Signals, Systems, and Transforms*, Addison-Wesley, Reading, MA, 1991.

Oppenheim, A.V., A.S. Willsky, S.H. Nawab, *Signal and Systems*, Prentice-Hall, Upper Saddle River, NJ, 1997.

Poularikas, A.D., S. Seely, *Signals and Systems*, PWS, Boston, 1985.

Appendix E

Answers to Selected Problems

Note: For problems based on computer simulations the number of realizations used in the computer simulation will affect the numerical results. In the results listed below the number of realizations is denoted by N_{real}. Also, each result assumes that `rand('state',0)` and/or `randn('state',0)` have been used to initialize the random number generator (see Appendix 2A for further details).

Chapter 1

1. experiment: toss a coin; outcomes: {head, tail}; probabilities: $1/2, 1/2$

5. a. continuous; b. discrete; c. discrete; d. continuous; e. discrete

7. yes, yes

10. $P[k = 9] = 0.0537$, probably not

13. $1/2$

14. 0.9973 for $\Delta = 0.001$

Chapter 2

1. $\hat{P}[Y = 0] = 0.7490$, $\hat{P}[Y = 1] = 0.2510$ ($N_{\text{real}} = 1000$)

3. via simulation: $\hat{P}[-1 \leq X \leq 1] = 0.6863$; via numerical integration with $\Delta = 0.01$, $P[-1 \leq X \leq 1] = 0.6851$ ($N_{\text{real}} = 10,000$)

6. values near zero

8. estimated mean $= 0.5021$; true mean $= 1/2$ ($N_{\text{real}} = 1000$)

11. estimated mean $= 1.0042$; true mean $= 1$ ($N_{\text{real}} = 1000$)

13. 1.2381 ($N_{\text{real}} = 1000$)

14. no; via simulation: mean of $\sqrt{U} = 0.6589$; via simulation: $\sqrt{\text{mean of } U} = 0.7125$ ($N_{\text{real}} = 1000$)

Chapter 3

1. a. $A^c = \{x : x \leq 1\}$, $B^c = \{x : x > 2\}$
b. $A \cup B = \{x : -\infty < x < \infty\} = \mathcal{S}$, $A \cap B = \{x : 1 < x < 2\}$
c. $A - B = \{x : x > 2\}$, $B - A = \{x : x \leq 1\}$

7. $A = \{1, 2, 3\}$, $B = \{4, 5\}$, $C = \{1, 2, 3\}$, $D = \{4, 5, 6\}$

10. $A \cup B \cup C = (A^c \cap B^c \cap C^c)^c$, $A \cap B \cap C = (A^c \cup B^c \cup C^c)^c$

12. a. 10^7, discrete b. 1, discrete c. ∞ (uncountable), continuous d. ∞ (uncountable), continuous e. 2, discrete f. ∞ (countable), discrete

14. a. $\mathcal{S} = \{t : 30 \leq t \leq 100\}$ b. outcomes are all t in interval $[30, 100]$ c. set of outcomes having no elements, i.e., {negative temperatures} d. $A = \{t : 40 \leq t \leq 60\}$, $B = \{t : 40 \leq t \leq 50 \text{ or } 60 \leq t \leq 70\}$, $C = \{100\}$ (simple event) e. $A = \{t : 40 \leq t \leq 60\}$, $B = \{t : 60 \leq t \leq 70\}$

18. a. 1/2 b. 1/2 c. 6/36 d. 24/36

19. $P_{\text{even}} = 1/2$, $\hat{P}_{\text{even}} = 0.5080$ ($N_{\text{real}} = 1000$)

21. a. even, 2/3 b. odd, 1/3 c. even or odd, 1 d. even and odd, 0

23. 1/56

25. 10/36

27. no

33. 90/216

35. 676,000

38. 0.00183

40. total number = 16, two-toppings = 6

44. a. 4 of a kind

$$\frac{13 \cdot 48}{\binom{52}{5}}$$

b. flush

$$\frac{4 \cdot \binom{13}{5}}{\binom{52}{5}}$$

49. $P[k \geq 95] = 0.4407$, $\hat{P}[k \geq 95] = 0.4430$ ($N_{\text{real}} = 1000$)

Chapter 4

2. 1/4

5. 1/4

7. a. 0.53 b. 0.34

11. 0.5

14. yes

19. 0.03

21. a. no b. no

22. 4

26. 0.0439

28. 5/16

33. $P[k] = (k-1)(1-p)^{k-2}p^2$, $k = 2, 3, \ldots,$

38. 2 red, 2 black, 2 white

40. 3/64

43. 165/512

Chapter 5

4. $\mathcal{S}_X = \{0, 1, 4, 9\}$

$$p_X[x_i] = \begin{cases} \frac{1}{7} & x_i = 0 \\ \frac{2}{7} & x_i = 1 \\ \frac{2}{7} & x_i = 4 \\ \frac{2}{7} & x_i = 9 \end{cases}$$

6. $0 < p < 1$, $\alpha = (1-p)/p^2$

8. 0.99^{19}

13. Average value = 5.0310, true value shown in Chapter 6 to be $\lambda = 5$ ($N_{\text{real}} = 1000$)

14. $p_X[5] = 0.0029$, $\hat{p}_X[5] = 0.0031$ (from Poisson approximation)

18. $P[X=3] = 0.0613$, $\hat{P}[X=3] = 0.0607$ ($N_{\text{real}} = 10,000$)

20. $p_Y[k] = \exp(-\lambda)\lambda^{k/2}/k!$ for $k = 0, 2, 4, \ldots$

26. $p_X[k] = 1/5$ for $k = 1, 2, 3, 4, 5$

28. 0.4375

31. 8.68×10^{-7}

Chapter 6

2. 9/2

4. 2/3

8. geometric PMF

12. $(2/p^2) - 1/p$

13. yes, if X = constant

14. predictor = $E[X] = 21/8$, $\text{mse}_{\min} = 47/64 = 0.7343$

15. estimated $\text{mse}_{\min} = 0.7371$ ($N_{\text{real}} = 10,000$)

20. $\lambda^2 + \lambda$

26. $\sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} E^{n-k}[X] E[X^k]$

27. $\phi_Y(\omega) = \exp(j\omega b)\phi_X(a\omega)$

28. $(1 + 2\cos(\omega) + 2\cos(2\omega))/5$

32. true mean = 1/2, true variance = 3/4 ; estimated mean = 0.5000, estimated variance = 0.7500 ($N_{\text{real}} = 1000$)

Chapter 7

3. $\mathcal{S} = \{(\text{p,n}), (\text{p,d}), (\text{n,p}), (\text{n,d}), (\text{d,p}), (\text{d,n})\}$
$\mathcal{S}_{X,Y} = \{(1,5), (1,10), (5,1), (5,10), (10,1), (10,5)\}$

8.

$$p_{X,Y}[i,j] = \begin{cases} 1/4 & (i,j) = (0,0) \\ 1/4 & (i,j) = (1,-1) \\ 1/4 & (i,j) = (1,1) \\ 1/4 & (i,j) = (2,0) \end{cases}$$

10. 1/5

13. $p_X[i] = (1-p)^{i-1}p$ for $i = 1, 2, \ldots$ and same for $p_Y[j]$

16. $p_{X,Y}[0,0] = 1/4$, $p_{X,Y}[0,1] = 0$, $p_{X,Y}[1,0] = 1/8$, $p_{X,Y}[1,1] = 5/8$

19. no

23. yes, $X \sim \text{bin}(10, 1/2)$, $Y \sim \text{bin}(11, 1/2)$

27. $p_Z[0] = 1/4$, $p_Z[1] = 1/2$, $p_Z[2] = 1/4$, variance always increases when uncorrelated random variables are added

33. 1/8

37. 0

38. 3/22

40. minimum MSE prediction $= E_Y[Y] = 5/8$ and minimum MSE $= \text{var}(Y) = 15/64$ for no knowledge
minimum MSE prediction $= \hat{Y} = -(1/15)x + 2/3$ and minimum MSE $= \text{var}(Y)(1-\rho^2_{X,Y}) = 7/30$ based on observing outcome of X

41. $\hat{W} = 5.4109h - 205.0344$

43. $\rho_{W,Z} = \sqrt{\eta/(\eta+1)}$, where $\eta = E_X[X^2]/E_N[N^2]$

46. see solution for Problem 7.27

48. $p_{X,Y}[0,0] = 0.1190$, $p_{X,Y}[0,1] = 0.1310$, $p_{X,Y}[1,0] = 0.2410$, $p_{X,Y}[1,1] = 0.5090$ ($N_{\text{real}} = 1000$)

49. $\rho_{X,Y} = \sqrt{5}/15 = 0.1490$, $\hat{\rho}_{X,Y} = 0.1497$ ($N_{\text{real}} = 100,000$)

Chapter 8

2. $p_{Y|X}[j|0] = 1$ for $j = 0$
$p_{Y|X}[j|1] = 1/6$ for $j = 1, 2, 3, 4, 5, 6$
$P[Y = 1] = 1/12$

5. no, no, no

6. $p_{Y|X}[j|0] = 1/3$ for $j = 0$ and $= 2/3$ for $j = 1$
$p_{Y|X}[j|1] = 2/3$ for $j = 0$ and $= 1/3$ for $j = 1$
$p_{X|Y}[i|0] = 1/3$ for $i = 0$ and $= 2/3$ for $i = 1$
$p_{X|Y}[i|1] = 2/3$ for $i = 0$ and $= 1/3$ for $i = 1$

8. $p_{Y|X}[j|i] = 1/5$ for $j = 0, 1, 2, 3, 4; i = 1, 2$
$p_{X|Y}[i|j] = 1/2$ for $i = 1, 2; j = 0, 1, 2, 3, 4$

11. 0.4535

13. a. $p_{Y|X}[y_j|0] = 0, 1, 0$ for $y_j = -1/\sqrt{2}, 0, 1/\sqrt{2}$, respectively
$p_{Y|X}[y_j|1/\sqrt{2}] = 1/2, 0, 1/2$ for $y_j = -1/\sqrt{2}, 0, 1/\sqrt{2}$, respectively
$p_{Y|X}[y_j|\sqrt{2}] = 0, 1, 0$ for $y_j = -1/\sqrt{2}, 0, 1/\sqrt{2}$, respectively
not independent (conditional PMF depends on x_i)
b. $p_{Y|X}[y_j|0] = 1/2, 1/2$ for $y_j = 0, 1$, respectively
$p_{Y|X}[y_j|1] = 1/2, 1/2$ for $y_j = 0, 1$, respectively
independent

17. $p_Z[k] = p_X[k]\sum_{j=k}^{\infty} p_Y[j] + p_Y[k]\sum_{j=k+1}^{\infty} p_X[j]$

21. $E_{Y|X}[Y|0] = 0$, $E_{Y|X}[Y|1] = 1/2$, $E_{Y|X}[Y|2] = 1$

22. $\text{var}(Y|0) = 0$, $\text{var}(Y|1) = 1/4$, $\text{var}(Y|2) = 2/3$

28. optimal predictor: $\hat{Y} = 0$ for $x = -1$, $\hat{Y} = 1/2$ for $x = 0$, and $\hat{Y} = 0$ for $x = 1$
optimal linear predictor: $\hat{Y} = 1/4$ for $x = -1, 0, 1$

30. $\widehat{E_{Y|X}}[Y|0] = 0.5204$, $\widehat{E_{Y|X}}[Y|1] = 0.6677$ ($N_{\text{real}} = 10{,}000$)

Chapter 9

1. 0.0567

4. yes

6. (X_1, X_2) independent of X_3

10. $E[\bar{X}] = E_X[X]$, $\text{var}(\bar{X}) = \text{var}(X)/N$

13. $\mathbf{C}_X = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $\det(\mathbf{C}_X) = 0$, no

17. a. no, b. no , c. yes, d. no

20. $\mathbf{C}_X = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix}$

26. $\mathbf{A} = \begin{bmatrix} 0.9056 & 0.4242 \\ -0.4242 & 0.9056 \end{bmatrix}$ for MATLAB 5.2
$\mathbf{A} = \begin{bmatrix} -0.9056 & 0.4242 \\ 0.4242 & 0.9056 \end{bmatrix}$ for MATLAB 6.5, R13
$\text{var}(Y_1) = 7.1898$, $\text{var}(Y_2) = 22.8102$

35. $\mathbf{B} = \begin{bmatrix} \sqrt{3/2} & \sqrt{5/2} \\ -\sqrt{3/2} & \sqrt{5/2} \end{bmatrix}$

36. $\hat{\mathbf{C}}_X = \begin{bmatrix} 4.0693 & 0.9996 \\ 0.9996 & 3.9300 \end{bmatrix}$ $(N_{\text{real}} = 1000)$

Chapter 10

2. 1/80

4. a. no b. yes c. no

6. $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, and $\alpha_1 + \alpha_2 = 1$

12. 0.0252

14. Gaussian: 0.0013 Laplacian: 0.0072

17. first person probability = 0.393, first two persons probability = 0.090

19. $F_X(x) = 1/2 + (1/\pi)\arctan(x)$

22. $F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

28. 2.28%

30. eastern U.S.

33. yes

36. $c \approx 14$

40.

$$p_Y(y) = \begin{cases} \frac{\lambda}{4(y-1)^{3/4}} \exp[-\lambda(y-1)^{1/4}] & y \geq 1 \\ 0 & y < 1 \end{cases}$$

43. $p_Y(y) = p_X(y) + p_X(-y)$

46.

$$p_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

51.

$$\begin{aligned} P[-2 \leq X \leq 2] &= 1 - \tfrac{1}{2}\exp(-2) \\ P[-1 \leq X \leq 1] &= 1 - \tfrac{1}{2}\exp(-1) \\ P[-1 < X \leq 1] &= \tfrac{3}{4} - \tfrac{1}{2}\exp(-1) \\ P[-1 < X < 1] &= \tfrac{1}{2} - \tfrac{1}{2}\exp(-1) \\ P[-1 \leq X < 1] &= \tfrac{3}{4} - \tfrac{1}{2}\exp(-1) \end{aligned}$$

54. $g(U) = \sqrt{2\ln(1/(1-U))}$

Chapter 11

1. 7/6

10. ± 9.12

11. 0.1353

14. N

19. 0.0078

21. $\sqrt{E[U]} = \sqrt{1/2}$, $E[\sqrt{U}] = 2/3$

22. $E[s(t_0)] = 0$, $E[s^2(t_0)] = 1/2$

26. $\sigma^2/2$

27. $\sigma^2/2$

30. $T_{\min} = 5.04$, $T_{\max} = 8.96$

35. $E[X^3] = 3\mu\sigma^2 + \mu^3$, $E[(X-\mu)^3] = 0$

38. $E[X^n] = 0$ for n odd, $E[X^n] = n!$ for n even

42. $\delta(x-\mu)$

44. $\sqrt{2\text{var}(X)}$

46. $E[X] = 1.2533$, $\widehat{E[X]} = 1.2538$; $\text{var}(X) = 0.4292$, $\widehat{\text{var}(X)} = 0.4269$ ($N_{\text{real}} = 1000$)

Chapter 12

1. 7/16

3. no, probability is 1/4

5. $\pi = 4P[X^2 + Y^2 \leq 1]$, $\hat{\pi} = 3.1140$ ($N_{\text{real}} = 10,000$)

7. 1/4

10. $P = 0.19$, $\hat{P} = 0.1872$ ($N_{\text{real}} = 10,000$)

11. 0

15. $p_X(x) = 2x$ for $0 < x < 1$ and zero otherwise, $p_Y(y) = 2(1-y)$ for $0 < y < 1$ and zero otherwise

18.

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ \frac{1}{8}xy & 0 \le x < 2, 0 \le y < 4 \\ \frac{1}{4}y & x \ge 2, 0 \le y < 4 \\ \frac{1}{2}x & 0 \le x < 2, y \ge 4 \\ 1 & x \ge 2, y \ge 4 \end{cases}$$

23. $(1 - \exp(-2))^2$

25. no

26. $Q(2)$

30. $P[\text{bullseye}] = 1 - \exp(-2) = 0.8646$, $\hat{P}[\text{bullseye}] = 0.8730$ $(N_{\text{real}} = 1000)$

36. $W \sim \mathcal{N}(\mu_W, \sigma_W^2)$, $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$

38. $\begin{bmatrix} W \\ Z \end{bmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$, where

$$\begin{aligned} \boldsymbol{\mu} &= \begin{bmatrix} 3 \\ 8 \end{bmatrix} \\ \mathbf{C} &= \begin{bmatrix} 2 & 5 \\ 5 & 14 \end{bmatrix} \end{aligned}$$

43. $\sqrt{5\pi}$

45. uncorrelated but not necessarily independent

47.

$$\mathbf{G} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

52. $Q(1)$

Chapter 13

2. yes, $c = 1/x$

4. $p_{Y|X}(y|x) = \exp(-y)/(1 - \exp(-x))$ for $0 \le y \le x$, $x \ge 0$

8. $p_{X,Y}(x,y) = 1/x$ for $0 < y < x$, $0 < x < 1$; $p_Y(y) = -\ln y$ for $0 < y < 1$

10. $p_{Y|X}(y|x) = 1/x$ for $0 < y < x$, $0 < x < 1$; $p_{X|Y}(x|y) = 1/(1-y)$ for $y < x < 1$, $0 < y < 1$

14. Use $P = \int_{-\frac{1}{2}}^{\frac{1}{2}} P[|X_2| - |X_1| < 0 | X_1 = x_1] p_{X_1}(x_1) dx_1$ and note independence of X_1 and X_2 so that $P = \int_{-\frac{1}{2}}^{\frac{1}{2}} P[|X_2| \leq x_1] dx_1$

16. $Q(-1)$, assume R and E are independent

21. $1/2$

24. Use $E_{(X+Y)|X}[X+Y|x] = E_{Y|X}[Y|x] + x$ to yield $E_{(X+Y)|X}[X+Y|X=50] = 77.45$ and $E_{(X+Y)|X}[X+Y|X=75] = 84.57$

Chapter 14

1. $E_Y[Y] = 6$, $\text{var}(Y) = 11/2$

6. $1/16$

9. $Y \sim \mathcal{N}(0, \sigma_1^2 + \sigma_2^2 + \sigma_3^2)$

12. no since $\text{var}(\bar{X}) \to \sigma^2/2$ as $N \to \infty$

19. $E_Y[Y] = 0$, $\text{var}(Y) = 1$

21. $\hat{X}_3 = 7/5$

24. $\text{mse}_{\text{min}} = 8/15 = 0.5333$

25. $\widehat{\text{mse}}_{\text{min}} = 0.5407$ ($N_{\text{real}} = 5000$)

Chapter 15

4. $\lim_{N\to\infty} \sum_{i=1}^{N} \alpha_i^2 = 0$, $\alpha_i = 1/N^{3/4}$

7. no since the variance does not converge to zero

13. $Y \sim \mathcal{N}(2000, 1000/3)$

19. $N = 5529$

20. $1 - Q(-77.78) \approx 0$

22. Gaussian, "converges" for all $N \geq 1$

23. no since approximate 95% confidence interval is $[0.723, 0.777]$

26. drug group has approximate 95% confidence interval of $[0.69, 0.91]$ and placebo group has $[0.47, 0.73]$. Can't say if drug is effective since true value of p could be 0.7 for either group.

Chapter 16

1. a. temperature at noon b. expense in dollars and cents c. time left in hours and minutes

4. $p^{50}(1-p)^{50}$, 0

7. $E[X[n]] = (n+1)/2$, $\text{var}(X[n]) = (3/4)(n+1)$

9. $\exp(-3)$

13. independent but not stationary

16. $\mu_X[n] = 0$, $c_X[n_1, n_2] = \delta[n_2 - n_1]$, exactly the same as for WGN with $\sigma_U^2 = 1$

18. $P[X[n] > 3] = 0.000011$, $P[U[n] > 3] = 0.0013$

22. $E[X(t)] = 0$, $c_X(t_1, t_2) = \cos(2\pi t_1)\cos(2\pi t_2)$

24. $E[Y[n]] = 0$, $\text{cov}(Y[0], Y[1]) = -1$, not IID since samples are not independent

26. $E[X[n]] = 0$, $c_X[n_1, n_2] = \sigma_U^2 \min(n_1, n_2)$

27. $c_X[1,1] = 1/2$, $c_X[1,2] = 1/4$, $c_X[1,3] = 0$, $\hat{c}_X[1,1] = 0.5057$, $\hat{c}_X[1,2] = 0.2595$, $\hat{c}_X[1,3] = -0.0016$ ($N_{\text{real}} = 10,000$)

31. $\mu_X[n] = 0$, $c_X[n_1, n_2] = \sigma_A^2 \sigma_U^2 \delta[n_2 - n_1]$, white noise

34. $N = 209$

Chapter 17

1. yes, $\mu_X[n] = \mu = 2p - 1$, $r_X[k] = 1$ for $k = 0$ and $r_X[k] = \mu^2$ for $k \neq 0$

5. WSS but not stationary since $p_{X[0]} \neq p_{X[1]}$

9. $a > 0$, $|b| \leq 1$

12. b,d,e

17. $E[X[n]] = 0$, $\text{var}(X[n]) = \sigma_U^2(1 - a^{2(n+1)})/(1 - a^2)$; as $n \to \infty$, $\text{var}(X[n]) \to \sigma_u^2/(1 - a^2)$

19. Principal minors are 1, 15/64, and $-17/32$ for 1×1, 2×2 and 3×3, respectively. Fourier transform is $1 - (7/4)\cos(2\pi f)$ which can be negative.

20. $\mu_X[n] = \mu$, $r_X[k] = (1/2)r_U[k] + (1/4)r_U[k-1] + (1/4)r_U[k+1]$

28. $P_X(f) = 2\sigma_U^2(1 - \cos(2\pi f))$

30. $P_X(f) = \sigma_A^2\sigma_U^2$

34. $r_X[k] = \sigma_U^2\delta[k] + \mu^2$, $P_X(f) = \sigma_U^2 + \mu^2\delta(f)$

38. $r_X[k] = 9/4, 3/2, 1/2$ for $k = 0, 1, 2$, respectively, and zero otherwise

40. $a \geq |b|$

42. $E[X[n]] = 0$, $\widehat{E}[X[10]] = -0.0105$, $\widehat{E}[X[12]] = 0.0177$; $r_X[2] = 0.1545$, $\widehat{E}[X[10]X[12]] = 0.1501$, $\widehat{E}[X[12]X[14]] = 0.1533$ ($N_{\text{real}} = 1000$)

44. $P_X(f) = 1$, increasing N does not improve estimate – must average over ensemble of periodograms

47. $2(\exp(-10) - \exp(-100))$

50. $\mu_X(t) = 0$, $\text{var}(X(t)) = N_0/(2T)$, no

51. $\text{var}(\hat{\mu}_N) = (1/N)\sum_{k=-(N-1)}^{N-1}(1 - |k|/N)N_0W\sin(\pi k/2)/(\pi k/2)$ for $N = 20$. It is 0.9841 times that of the variance of the sample mean for Nyquist rate sampled data.

Chapter 18

1. $r_X[k] = 3$ for $k = 0$, $r_X[k] = -1$ for $k = \pm 2$, and equals zero otherwise; $P_X(f) = 3 - 2\cos(4\pi f)$

4. $b_1 = 0$, $b_2 = -1$

7. $r_X[k] = 3$ for $k = 0$, $r_X[k] = -2$ for $k = \pm 1$, $r_X[k] = 1/2$ for $k = \pm 2$, and equals zero otherwise; $P_X(f) = 3 - 4\cos(2\pi f) + \cos(4\pi f)$

13. $H_{\text{opt}} = (2 - 2\cos(2\pi f))/(3 - 2\cos(2\pi f))$; $\text{mse}_{\min} = 0.5552$

18. $\text{mse}_{\min} = r_X[0](1 - \rho^2_{X[n_0],X[n_0+1]})$

22. $\hat{X}[n_0 + 1] = -[b(1 + b^2)/(1 + b^2 + b^4)]X[n_0] - [b^2/(1 + b^2 + b^4)]X[n_0 - 1]$; $\text{mse}_{\min} = 1 + b^6/(1 + b^2 + b^4)$

24. $\text{mse}_{\min} = 1 + [b^6/(1 + b^2 + b^4)] = 85/84 = 1.0119$, $\widehat{\text{mse}}_{\min} = 1.0117$ ($N_{\text{real}} = 10,000$)

27. $\hat{X}[n_0] = [a/(1 + a^2)](X[n_0 + 1] + X[n_0 - 1])$

29. $P_X(F) = (N_0T^2/2)[(\sin(\pi FT))/(\pi FT)]^2$

32. $r_X(0) = N_0/(4RC)$, no

Chapter 19

1. $E[X[n]Y[n+k]] = (-1)^{n+k}\sigma_U^2\delta[k]$, no

5. $r_{X,Y}[k] = 0$

6. $r_{X,U}[k] = 0$ for $k > 0$ and $r_{X,Y}[k] = (1/2)^{(-k)}$ for $k < 0$

10. $|P_{X,Y}(f)| = \sqrt{5 + 4\cos(2\pi f)}$, $\angle P_{X,Y}(f) = \arctan \frac{-2\sin(2\pi f)}{1+2\cos(2\pi f)}$

12. $r_Z[k] = r_X[k] - r_{X,Y}[k] - r_{Y,X}[k] + r_Y[k]$, $P_Z(f) = P_X(f) - P_{X,Y}(f) - P_{Y,X}(f) + P_Y(f)$

15. $\gamma_{X,Y}(f) = -1$, perfectly predictable using $\hat{Y}[n_0] = -X[n_0]$

18. $H_{\text{opt}}(f) = P_{X,Y}(f)/P_X(f)$

23. $r_{X,Y}(\tau) = N_0/(2T)$ for $0 \le \tau \le T$ and zero otherwise

26. $r_{X,Y}[k] = \delta[k] - b\delta[k-1]$, for $b = -1$

k	$r_{X,Y}[k]$	$\hat{r}_{X,Y}[k]$	k	$r_{X,Y}[k]$	$\hat{r}_{X,Y}[k]$
−5	0	−0.0077	0	1	0.9034
−4	0	−0.0242	1	1	0.9031
−3	0	0.0259	2	0	−0.0064
−2	0	0.0004	3	0	−0.0007
−1	0	−0.0062	4	0	0.0267
			5	0	−0.0238

Chapter 20

2. 1/4

5. $\mathbf{Y} = [Y[0]\,Y[1]]^T \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_Y)$, where

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

not independent

10. WSS with $\mu_Z = \mu_X\mu_Y$, $P_Z(f) = P_X(f) \star P_Y(f)$

14. 1

17. $T = 66,347$

19. $2Q(1/\sqrt{T})$

22. $\mathbf{Y} = [Y(0)\,Y(1/4)]^T \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_Y)$, where

$$\mathbf{C}_Y = \begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 2 \end{bmatrix}$$

25. $P_U(F) = P_V(F) = 8(1 - |F|/10)$ for $|F| \leq 10$ and zero otherwise

30. $X[n] = U[n] - U[n-1]$, where $U[n]$ is WGN with $\sigma_U^2 = 1$

31. $r_X[0] = 2$, $\hat{r}_X[0] = 1.9591$; $r_X[1] = -1$, $\hat{r}_X[1] = -0.9614$; $r_X[2] = 0$, $\hat{r}_X[2] = -0.0195$; $r_X[3] = 0$, $\hat{r}_X[3] = -0.0154$

Chapter 21

3. probability $= 0.1462$, average $= 5$

7. $\lambda = 2, \hat{\lambda} = 1.9629$; $\lambda = 5, \hat{\lambda} = 4.9072$ (based on 10,000 arrivals with $\hat{\lambda} = \hat{E}[N(t)]/t$)

10. $E[N(t_2) - N(t_1)] = \text{var}(N(t_2) - N(t_1)) = \lambda(t_2 - t_1)$,

13. 0.6321

17. 10 minutes

20. $P[T_2 \leq 1] = 1 - 2\exp(-1) = 0.2642$, $\hat{P}[T_2 \leq 1] = 0.2622$ ($N_{\text{real}} = 10,000$)

23. $\lambda t_0(2p-1)$

Chapter 22

2. 1/128

5. $P[Y[2] = 1|Y[1] = 1, Y[0] = 0] = 1 - p$, $P[Y[2] = 1|Y[1] = 1] = 1/2$ for all p

9. $P_e = 0.3362$

11. $P[\text{red drawn}] = 1/3$

12. 1/2

14. yes, $\boldsymbol{\pi}^T = [\frac{1}{3}\ \frac{2}{3}\ 0\ 0]$

19. $\boldsymbol{\pi}^T = [0.2165\ 0.4021\ 0.3814]$

24. $P[\text{rain}] = 0.6964$

26. $n = 6$

28. $\pi_1 = 1/3$, $\hat{\pi}_1 = 0.3240$ (based on playing 1000 holes)

Index