“Signal Processing: A Mathematical Approach” - Answers to Selected Exercises

Charles L. Byrne

March 26, 2008

Chapter 1: Complex Numbers

Exercise 1.1: Derive the formula for dividing one complex number in rectangular form by another (non-zero) one.

Solution: For any complex numbers $z = (a, b)$ its reciprocal $z^{-1} = (c, d)$ must satisfy the equation $zz^{-1} = (1, 0) = 1$. Therefore $ac - bd = 1$ and $ad + bc = 0$. Multiplying the first equation by a and the second by b and adding, we get $(a^2 + b^2)c = a$, so $c = a/(a^2 + b^2)$. Inserting this in place of c in the second equation gives $d = -b/(a^2 + b^2)$. To divide any complex number w by z we multiply w by z^{-1}.

Exercise 1.2: Show that for any two complex numbers z and w we have

$$|zw| \geq \frac{1}{2}(z\overline{w} + \overline{z}w).$$

Hint: Write $|zw|$ as $|z\overline{w}|$.

Solution: Using the polar form for z and w it is easy to see that $|zw| = |z\overline{w}|$. With $v = z\overline{w}$ the problem is now to show that $|v| \geq \frac{1}{2}(v + \overline{v})$, or $|v| \geq Re(v)$, which is obvious.

Chapter 2: Complex Exponentials

Exercise 2.2: The *Dirichlet kernel* of size M is defined as

$$D_M(x) = \sum\nolimits_{m=-M}^{M} e^{imx}.$$

Obtain the closed-form expression

$$D_M(x) = \frac{\sin((M + \frac{1}{2})x)}{\sin(\frac{x}{2})};$$

note that $D_M(x)$ is real-valued.
Hint: Reduce the problem to that of Exercise 2.1 by factoring appropriately.

Solution: Factor out the term $e^{-i(M+1)x}$ to get

$$D_M(x) = e^{-i(M+1)x} \sum_{m=1}^{2M+1} e^{imx}.$$

Now use the solution to the previous exercise.

Exercise 2.3: Use the formula for $E_M(x)$ to obtain the closed-form expressions

$$\sum_{m=N}^{M} \cos mx = \cos(\frac{M+N}{2}x)\frac{\sin(\frac{M-N+1}{2}x)}{\sin\frac{x}{2}}$$

and

$$\sum_{m=N}^{M} \sin mx = \sin(\frac{M+N}{2}x)\frac{\sin(\frac{M-N+1}{2}x)}{\sin\frac{x}{2}}.$$

Hint: Recall that $\cos mx$ and $\sin mx$ are the real and imaginary parts of e^{imx}.

Solution: Begin with

$$S(x) = \sum_{m=N}^{M} e^{imx}$$

and factor out $e^{i(N-1)x}$ to get

$$S(x) = e^{i(N-1)x} \sum_{m=1}^{M-N+1} e^{imx}.$$

Now apply the formula for $E_M(x)$. Finally, use the fact that the two sums we seek are the real and imaginary parts of $S(x)$.

Chapter 3: Hidden Periodicities

Exercise 3.1: Determine the formulas giving the horizontal and vertical coordinates of the position of a particular rider at an arbitrary time t in the time interval $[0, T]$.

Solution: Since the choice of the origin of our coordinate system is arbitrary, we take the origin $(0, 0)$ to be the point on the ground directly under the center of the wheel. The center of the wheel is then located at the point $(0, R + H)$. Let the rider be at the point $(0 + R\cos\theta, R + H + R\sin\theta)$ at time $t = 0$. Since the wheel turns with angular frequency ω the horizontal position of the rider at any subsequent time will be

$$x(t) = 0 + R\cos(\theta + t\omega)$$

and the vertical position will be

$$y(t) = R + H + R\sin(\theta + t\omega).$$

Note that we can represent the rider's position as a complex number

$$0 + (R + H)i + Re^{i(\theta + t\omega)}.$$

Exercise 3.2: Now find the formulas giving the horizontal and vertical coordinates of the position of a particular rider at an arbitrary time t in the time interval $[0, T]$.

Solution: The position of the center of the smaller wheel is the same as that of the rider in the previous exercise; that is,

$$x(t) = 0 + R_1 \cos(\theta_1 + t\omega_1)$$

and

$$y(t) = R_1 + H + R_1 \sin(\theta_1 + t\omega_1).$$

The rider's position deviates from that of the center of the smaller wheel in the same way that the rider's position in the previous exercise deviated from the center of the single large wheel. Therefore, the horizontal position of the rider now is

$$x(t) = 0 + R_1 \cos(\theta_1 + t\omega_1) + R_2 \cos(\theta_2 + t\omega_2)$$

and the vertical position is

$$y(t) = R_1 + H + R_1 \sin(\theta_1 + t\omega_1) + R_2 \sin(\theta_2 + t\omega_2).$$

Again, we can represent the position as a complex number:

$$0 + (R + H)i + R_1 e^{i(\theta_1 + t\omega_1)} + R_2 e^{i(\theta_2 + t\omega_2)}.$$

Exercise 3.3: Repeat the previous exercise, but for the case of J nested wheels.

Solution: Reasoning as above, and using the complex representation, we find the position to be

$$0 + (R + H)i + \sum_{j=1}^{J} R_j e^{i(\theta_j + t\omega_j)}.$$

Chapter Five: Convolution and the Discrete Fourier Transform

Exercise 5.1: Let $\mathbf{F} = vDFT_{\mathbf{f}}$ and $\mathbf{D} = vDFT_{\mathbf{d}}$. Define a third vector $\mathbf{E}$ having for its k-th entry $E_k = F_k D_k$, for $k = 0, ..., N - 1$. Show that $\mathbf{E}$ is the vDFT of the vector $\mathbf{f} * \mathbf{d}$.

Solution: For notational convenience we define $d_{k-N} = d_k$, for $k = 0, 1, ..., N$. Then we can write

$$(\mathbf{f} * \mathbf{d})_n = \sum_{m=0}^{N-1} f_m d_{n-m}.$$

Using this extended notation we find that the sum

$$\sum_{n=0}^{N-1} d_{n-m} e^{i(n-m)2\pi k/N}$$

does not depend on m and is equal to

$$\sum_{j=0}^{N-1} d_j e^{2\pi jki/N},$$

which is D_k. The vDFT of the vector $\mathbf{f} * \mathbf{d}$ has for its k-th entry the quantity

$$\sum_{n=0}^{N-1} (\mathbf{f} * \mathbf{d})_n e^{2\pi ink/N},$$

which we write as the double sum

$$\sum_{n=0}^{N-1} \sum_{m=0}^{N-1} f_m d_{n-m} e^{2\pi ink/N}.$$

Now we simply reverse the order of summation, write

$$e^{2\pi ink/N} = e^{2\pi imk/N} e^{2\pi i(n-m)k/N}$$

and use the fact already shown that the sum on n is independent of m. We then have that the k-th entry is

$$\sum_{m=0}^{N-1} f_m e^{2\pi imk/N} \sum_{j=0}^{N-1} d_j e^{2\pi ijk/N} = F_k D_k.$$

Exercise 5.2: Let G be the N by N matrix whose entries are $G_{jk} = e^{i(j-1)(k-1)2\pi/N}$. The matrix G is sometimes called the *DFT matrix*. Show that the inverse of G is $G^{-1} = \frac{1}{N} G^\dagger$, where $G^\dagger$ is the conjugate transpose of the matrix G. Then $\mathbf{f} * \mathbf{d} = G^{-1}\mathbf{E} = \frac{1}{N} G^\dagger \mathbf{E}$.

Solution: Compute the entry of the matrix $G^\dagger G$ in the m-th row, n-th column. Use the definition of matrix multiplication to express this entry as a sum of the same type as in the definition of $E_M(x)$. Consider what happens when $m = n$ and when $m \neq n$.

Chapter 6: Inner Products

Exercise 6.1: Use Cauchy's inequality to show that

$$||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||;$$

this is called the *triangle inequality.*

Solution: We have

$$||\mathbf{u}+\mathbf{v}||^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}$$

$$= ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + \mathbf{u}\cdot\mathbf{v} + \overline{\mathbf{u}\cdot\mathbf{v}} = ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + 2Re(\mathbf{u}\cdot\mathbf{v}).$$

Also we have

$$(||\mathbf{u}|| + ||\mathbf{v}||)^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + 2||\mathbf{u}||\,||\mathbf{v}||.$$

Now use Cauchy's inequality to conclude that

$$Re(\mathbf{u}\cdot\mathbf{v}) \leq |Re(\mathbf{u}\cdot\mathbf{v})| \leq |\mathbf{u}\cdot\mathbf{v}| \leq ||\mathbf{u}||\,||\mathbf{v}||.$$

Exercise 6.2: Use the Gram-Schmidt approach to find a third vector in R^3 orthogonal to both $(1,1,1)$ and $(1,0,-1)$.

Solution: Let the third vector be $\mathbf{v} = (a,b,c)$. Notice that any vector that can be written as $\alpha(1,1,1)+\beta(1,0,-1)$ must have the form $(\alpha+\beta,\alpha,\alpha-\beta)$, so that the second entry is the average of the first and third entries. Now take any vector that does not have this property; for example, let's take $(1,2,2)$. We know that we can write $(1,2,2)$ as

$$(1,2,2) = \alpha(1,1,1) + \beta(1,0,-1) + \gamma(a,b,c),$$

for some choices of α, β and γ. Let's find α and β. Take the dot product of both sides of the last equation with the vector $(1,1,1)$ to get

$$5 = (1,1,1)\cdot(1,2,2) = \alpha(1,1,1)\cdot(1,1,1) = 3\alpha.$$

So $\alpha = 5/3$. Now take the inner product of both sides with $(1,0,-1)$ to get

$$-1 = (1,0,-1)\cdot(1,2,2) = \beta(1,0,-1)\cdot(1,0,-1) = 2\beta.$$

Therefore, $\beta = -1/2$. So we now have

$$(1,2,2) - \frac{5}{3}(1,1,1) + \frac{1}{2}(1,0,-1) = (-\frac{1}{6},\frac{1}{3},-\frac{1}{6}) = \frac{-1}{6}(1,-2,1).$$

We can then take $\gamma = \frac{-1}{6}$ and $\mathbf{v} = (a,b,c) = (1,-2,1)$.

Exercise 6.3: Find polynomial functions $f(x)$, $g(x)$ and $h(x)$ that are orthogonal on the interval $[0,1]$ and have the property that every polynomial of degree two or less can be written as a linear combination of these three functions.

Solution: Let's find $f(x) = a$, $g(x) = bx + c$ and $h(x) = dx^2 + ex + k$ that do the job. Clearly, we can start by taking $f(x) = 1$. Then

$$0 = \int_0^1 1g(x)dx = b\int_0^1 xdx + c = \frac{b}{2} + c$$

says that $b = -2c$. Let $c = 1$ so that $b = -2$ and $g(x) = -2x + 1$. Then

$$0 = \int_0^1 1h(x)dx = \frac{d}{3} + \frac{e}{2} + k$$

and

$$0 = \int_0^1 g(x)h(x)dx = \int_0^1 (-2x+1)(dx^2 + ex + k)dx.$$

Therefore we have

$$0 = \frac{-2}{4}d + \frac{-2}{3}e + \frac{-2}{2}k + \frac{d}{3} + \frac{e}{2} + k.$$

We can let $d = 6$, from which it follows that $e = -6$ and $k = 1$. So the three polynomials are $f(x) = 1$, $g(x) = -2x + 1$ and $h(x) = 6x^2 - 6x + 1$. To show that any quadratic polynomial can be written as a sum of these three, take an arbitrary quadratic, $ax^2 + bx + c$ and write

$$ax^2 + bx + c = \alpha f(x) + \beta g(x) + \gamma h(x).$$

Then show that you can solve for the α, β and γ in terms of the a, b and c.

Exercise 6.4: Show that the functions e^{inx}, n an integer, are orthogonal on the interval $[-\pi, \pi]$. Let $f(x)$ have the Fourier expansion

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx},\ |x| \leq \pi.$$

Use orthogonality to find the coefficients a_n.

Solution: Compute the integral

$$\int_{-\pi}^{\pi} e^{inx}e^{-imx}dx$$

and show that it is zero for $m \neq n$. To find the coefficients multiply both sides by e^{-imx} and integrate; on the left we get $\int_{-\pi}^{\pi} f(x)e^{-imx}dx$ and on the right we get $2\pi a_m$.

Chapter 7: Discrete Linear Filters

Exercise 7.1: Show that $F(\omega) = G(\omega)H(\omega)$ for all ω.

Solution: Using the definition of $F(\omega)$ and f_n we write

$$F(\omega) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} g_m h_{n-m} e^{i\omega m} e^{i\omega(n-m)}$$

$$= \sum_{m=-\infty}^{\infty} g_m [\sum_{n=-\infty}^{\infty} h_{n-m} e^{i\omega(n-m)}] \, e^{i\omega m}.$$

Since the inner sum

$$\sum_{n=-\infty}^{\infty} h_{n-m} e^{i\omega(n-m)} = \sum_{k=-\infty}^{\infty} h_k e^{i\omega k}$$

does not really depend on the index m it can be taken outside the sum over that index.

Exercise 7.2: The *three-point moving average* filter is defined as follows: given the input sequence $\{h_n, n = -\infty, ..., \infty\}$ the output sequence is $\{f_n, n = -\infty, ..., \infty\}$, with

$$f_n = (h_{n-1} + h_n + h_{n+1})/3.$$

Let $g_k = 1/3$, if $k = 0, 1, -1$ and $g_k = 0$, otherwise. Then we have

$$f_n = \sum_{k=-\infty}^{\infty} g_k h_{n-k},$$

so that f is the *discrete convolution* of h and g. Let

$$F(\omega) = \sum_{n=-\infty}^{\infty} f_n e^{in\omega},$$

for ω in the interval $[-\pi, \pi]$, be the Fourier series for the sequence f; similarly define G and H. To recover h from f we might proceed as follows: calculate F, then divide F by G to get H, then compute h from H; does this always work? If we let h be the sequence $\{..., 1, 1, 1, ...\}$ then $f = h$; if we take h to be the sequence $\{..., 3, 0, 0, 3, 0, 0, ...\}$ then we again get $f = \{..., 1, 1, 1, ...\}$. Therefore, we cannot expect to recover h from f in general. We know that $G(\omega) = \frac{1}{3}(1 + 2\cos(\omega))$; what does this have to do with the problem of recovering h from f?

Solution: If the input sequence is $h = \{..., 2, -1, -1, 2, -1, -1, ...\}$ then the output sequence is $f = \{..., 0, 0, 0, 0, 0, ...\}$. Since

$$G(\omega) = \frac{1}{3}(1 + 2\cos(\omega)),$$

the zeros of $G(\omega)$ are at $\omega = \frac{2\pi}{3}$ and $\omega = -\frac{2\pi}{3}$. Consider the sequence defined by

$$h_n = e^{in\frac{2\pi}{3}} + e^{-in\frac{2\pi}{3}};$$

this is the sequence $\{..., 2, -1, -1, 2, -1, -1, ...\}$. This sequence consists of two complex exponential components, with associated frequencies at precisely the roots of $G(\omega)$. The three-point moving average has the output of all zeros because the function $G(\omega)$ has *nulled out* the only two sinusoidal components in h.

Exercise 7.3: Let f be the autocorrelation sequence for g. Show that $f_{-n} = \overline{f}_n$ and $f_0 \geq |f_n|$ for all n.

Solution: The first part follows immediately from the definition of the autocorrelation. The second part is a consequence of the Cauchy-Schwarz inequality for infinite sequences.

Exercise 7.7: Let $f(t) = e^{-i\omega t}$ for some fixed real number ω. Let $h = Tf$, where T is linear and time-invariant.'Show that there is a constant c so that $h(t) = cf(t)$. Since the constant c may depend on ω we rewrite c as $G(\omega)$.

Solution: Using the hint, we differentiate $h(t)$. Since T is time-invariant,

$$h(t + \Delta t) = h_{\Delta t}(t) = (Tf)_{\Delta t}(t) = T(f_{\Delta t})(t).$$

Since T is linear, and $f_{\Delta t}(s) = e^{-i\omega\Delta t}e^{-i\omega s}$, we have

$$T(f_{\Delta t})(t) = e^{-i\omega\Delta t}T(f)(t) = e^{-i\omega\Delta t}h(t).$$

Therefore,

$$\frac{h(t + \Delta t) - h(t)}{\Delta t} = h(t)\frac{e^{-i\omega\Delta t} - 1}{\Delta t};$$

the limit, as $\Delta t \to 0$, is $-i\omega h(t)$. Consequently, we know that $h'(t) = -i\omega h(t)$, from which it follows that $h(t) = ce^{-i\omega t} = cf(t)$, for some constant c.

Chapter 8: Fourier Transforms and Fourier Series

Exercise 8.1: Use the orthogonality of the functions $e^{im\omega}$ on $[-\pi, \pi]$ to establish *Parseval's equation*:

$$\langle f, g\rangle = \sum_{m=-\infty}^{\infty} f_m\overline{g_m} = \int_{-\pi}^{\pi} F(\omega)\overline{G(\omega)}d\omega/2\pi,$$

from which it follows that

$$\langle f, f\rangle = \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega/2\pi.$$

Solution: Since we have

$$F(\omega) = \sum\nolimits_{m=-\infty}^{\infty} f_m e^{im\omega},\ |\omega| \leq \pi,$$

with a similar expression for $G(\omega)$, we have

$$\langle F, G\rangle\ = \int_{-\pi}^{\pi} F(\omega)\overline{G(\omega)} d\omega/2\pi$$

$$= \int_{-\pi}^{\pi} \sum\nolimits_{m=-\infty}^{\infty} f_m e^{im\omega} \sum\nolimits_{n=-\infty}^{\infty} \overline{g_n} e^{-in\omega} d\omega/2\pi$$

$$= \sum\nolimits_{m=-\infty}^{\infty} \sum\nolimits_{n=-\infty}^{\infty} f_m \overline{g_n} \int_{-\pi}^{\pi} e^{i(n-m)\omega} d\omega/2\pi,$$

which equals

$$\sum\nolimits_{m=-\infty}^{\infty} f_m \overline{g_m} = \langle f, g\rangle$$

because the integral is zero unless $m = n$.

Exercise 8.2: Let $f(x)$ be defined for all real x and let $F(\omega)$ be its FT. Let

$$g(x) = \sum_{k=-\infty}^{\infty} f(x + 2\pi k),$$

assuming the sum exists. Show that g is a 2π -periodic function. Compute its Fourier series and use it to derive the *Poisson summation formula*:

$$\sum_{k=-\infty}^{\infty} f(2\pi k) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(n).$$

Solution: Clearly $g(x + 2\pi) = g(x)$ for all x, so $g(x)$ is 2π-periodic. The Fourier series for $g(x)$ is

$$g(x) = \sum\nolimits_{n=-\infty}^{\infty} a_n e^{inx},$$

where

$$a_n = \int_{-\pi}^{\pi} g(x) e^{-inx} dx/2\pi$$

$$= \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} f(x+2\pi k)e^{-inx}dx/2\pi$$

$$= \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} f(x+2\pi k)e^{-inx}dx/2\pi$$

$$= \sum_{k=-\infty}^{\infty} e^{i2\pi nk} \int_{-\pi}^{\pi} f(t)e^{-int}dt/2\pi$$

$$= \sum_{k=-\infty}^{\infty} \int_{-\pi}^{\pi} f(t)e^{-in(t-2\pi k)}dt/2\pi$$

$$= \sum_{k=-\infty}^{\infty} \int_{-\pi+2\pi k}^{\pi+2\pi k} f(t)e^{-int}dt/2\pi$$

$$= \int_{-\infty}^{\infty} f(t)e^{-int}dt/2\pi = \frac{1}{2\pi}F(-n).$$

Therefore

$$g(x) = \frac{1}{2\pi} \sum\nolimits_{n=-\infty}^{\infty} F(-n)e^{inx}.$$

Now let $x = 0$ to get

$$g(0) = \sum_{k=-\infty}^{\infty} f(2\pi k) = \frac{1}{2\pi} \sum\nolimits_{n=-\infty}^{\infty} F(-n).$$

Chapter 10: Fourier-Transform Pairs

Exercise 10.1: Let $F(\omega)$ be the FT of the function $f(x)$. Use the definitions of the FT and IFT to establish the following basic properties of the Fourier transform operation:

Differentiation: The FT of the n-th derivative, $f^{(n)}(x)$ is $(-i\omega)^n F(\omega)$. The IFT of $F^{(n)}(\omega)$ is $(ix)^n f(x)$.

Solution: Begin with the inverse FT equation

$$f(x) = \int F(\omega)e^{-ix\omega}d\omega/2\pi$$

and differentiate with respect to x inside the integral sign n times.

Convolution in x: Let f, F, g, G and h, H be FT pairs, with

$$h(x) = \int f(y)g(x-y)dy,$$

so that $h(x) = (f * g)(x)$ is the convolution of $f(x)$ and $g(x)$. Then $H(\omega) = F(\omega)G(\omega)$.

Solution: From the definitions of $F(\omega)$ and $G(\omega)$ we have

$$F(\omega)G(\omega) = \int f(y)e^{iy\omega}dy \int g(t)e^{it\omega}dt$$

$$= \int \int f(y)g(t)e^{i(y+t)\omega}dy\,dt.$$

Changing variables by setting $x = y + t$, so $t = x - y$ and $dt = dx$ we get

$$= \int \int f(y)g(x-y)e^{ix\omega}dydx$$

$$= \int [\int f(y)g(x-y)dy]\, e^{ix\omega}dx = \int h(x)e^{ix\omega}dx = H(\omega).$$

Exercise 10.2 Show that if T is a linear, time-invariant operator, then T is a convolution operator.

Solution: When the input is the function $f(t) = e^{-i\omega t}$, the output is $T(f)(t) = G(\omega)e^{-i\omega t}$. So with

$$f(t) = \frac{1}{2\pi}\int F(\omega)e^{-i\omega t}d\omega,$$

we have

$$T(f)(t) = \frac{1}{2\pi}\int F(\omega)G(\omega)e^{-i\omega t}d\omega.$$

So the Fourier transform of the function $T(f)(t)$ is $F(\omega)G(\omega)$, which tells us that $T(f)(t)$ is the convolution of $f(t)$ and $g(t)$.

Exercise 10.3: Show that the Fourier transform of $f(x) = e^{-\alpha^2 x^2}$ is $F(\omega) = \frac{\sqrt{\pi}}{\alpha}e^{-(\frac{\omega}{2\alpha})^2}$.

Solution: From the FT formula

$$F(\omega) = \int f(x)e^{ix\omega}dx = \int e^{-\alpha^2 x^2}e^{ix\omega}dx$$

we have

$$F'(\omega) = \int ixe^{-\alpha^2 x^2}e^{ix\omega}dx.$$

Integrating by parts gives

$$F'(\omega) = -\frac{\omega}{2\alpha^2}F(\omega),$$

so that

$$F(\omega) = c\exp(-\frac{\omega^2}{4\alpha^2}).$$

To find c we set $\omega = 0$. Then

$$c = F(0) = \int e^{-\alpha^2 x^2} dx = \frac{\sqrt{\pi}}{\alpha}.$$

This last integral occurs frequently in texts on probability theory, in the discussion of normal random variables and is obtained by using a trick involving polar coordinates. We calculate the square of this integral as

$$\int e^{-\alpha^2 x^2} dx \int e^{-\alpha^2 y^2} dy = \int\int e^{-\alpha^2(x^2+y^2)} dxdy = \int_0^{2\pi}\int_0^{\infty} e^{-\alpha^2 r^2} rdrd\theta$$

$$= \frac{2\pi}{2}\frac{-1}{\alpha^2} e^{-\alpha^2 r^2}|_0^{\infty} = \frac{2\pi}{2\alpha^2} = \frac{\pi}{\alpha^2}.$$

There is a second approach that we can use to solve this problem. We know that

$$f'(x) = -2\alpha^2 x f(x),$$

and that $ixf(x)$ is the inverse Fourier transform of $F'(\omega)$. Therefore, $f'(x)$ is the inverse Fourier transform of $2i\alpha^2 F'(\omega)$. But $f'(x)$ is also the inverse Fourier transform of $-i\omega F(\omega)$. It follows that

$$\omega F(\omega) + 2\alpha^2 F'(\omega) = 0.$$

We solve this first-order linear differential equation and proceed as above.

Exercise 10.4: Calculate the FT of the function $f(x) = u(x)e^{-ax}$, where a is a positive constant.

Solution: We have

$$F(\omega) = \int_0^{\infty} e^{-ax} e^{ix\omega} dx = \int_0^{\infty} e^{(i\omega - a)x} dx$$

$$= \frac{1}{i\omega - a}[\lim_{X\to+\infty}(e^{(i\omega-a)X}) - e^{(i\omega-a)(0)}] = \frac{1}{a - i\omega}.$$

Exercise 10.5: Calculate the FT of $f(x) = \chi_X(x)$.

Solution: We now have

$$F(\omega) = \int_{-X}^{X} e^{ix\omega} dx = \int_{-X}^{X} \cos(x\omega) dx$$

$$= \frac{2}{\omega}\sin(X\omega).$$

Exercise 10.7: Use the fact that $\operatorname{sgn}(x) = 2u(x) - 1$ and the previous exercise to show that $f(x) = u(x)$ has the FT $F(\omega) = i/\omega + \pi\delta(\omega)$.

Solution: From the previous exercise we know that the FT of $f(x) = \operatorname{sgn}(x)$ is $F(\omega) = \frac{2i}{\omega}$. We also know that the FT of the function $f(x) = 1$ is $F(\omega) = 2\pi\delta(\omega)$. Writing

$$u(x) = \frac{1}{2}(\operatorname{sgn}(x) + 1)$$

we find that the FT of $u(x)$ is $\frac{i}{\omega} + \pi\delta(\omega)$.

Exercise 10.8: Let $F(\omega) = R(\omega) + iX(\omega)$, where R and X are real-valued functions, and similarly, let $f(x) = f_1(x) + if_2(x)$, where f_1 and f_2 are real-valued. Find relationships between the pairs R,X and f_1,f_2.

Solution: From $F(\omega) = R(\omega) + iX(\omega)$ and

$$F(\omega) = \int f(x)e^{ix\omega}dx = \int (f_1(x) + if_2(x))e^{ix\omega}dx$$

we get

$$R(\omega) = \int f_1(x)\cos(x\omega) - f_2(x)\sin(x\omega)dx$$

and

$$X(\omega) = \int f_1(x)\sin(x\omega) + f_2(x)\cos(x\omega)dx.$$

Exercise 10.9: Let f, F be a FT pair. Let $g(x) = \int_{-\infty}^{x} f(y)dy$. Show that the FT of $g(x)$ is $G(\omega) = \pi F(0)\delta(\omega) + \frac{F(\omega)}{i\omega}$.

Solution: Since $g(x)$ is the convolution of $f(x)$ and the Heaviside function $u(x)$ it follows that

$$\begin{aligned} G(\omega) &= F(\omega)(\frac{i}{\omega} + \pi\delta(\omega)) \\ &= i\frac{F(\omega)}{\omega} + \pi F(0)\delta(\omega). \end{aligned}$$

Exercise 10.10: Let $f(x), F(\omega)$ and $g(x), G(\omega)$ be Fourier transform pairs. Establish the Parseval-Plancherel equation

$$\langle f, g\rangle = \int f(x)\overline{g(x)}dx = \frac{1}{2\pi}\int F(\omega)\overline{G(\omega)}d\omega.$$

Solution: Begin by inserting

$$f(x) = \int F(\omega)e^{-ix\omega}d\omega/2\pi$$

and

$$g(x) = \int G(\alpha)e^{-ix\alpha}d\alpha/2\pi$$

into

$$\int f(x)\overline{g(x)}dx$$

and interchanging the order of integration to get

$$\int f(x)\overline{g(x)}dx = (\frac{1}{2\pi})^2 \int \int F(\omega)\overline{G(\alpha)}[\int e^{ix(\omega-\alpha)}dx]d\omega d\alpha.$$

The innermost integral is

$$\int e^{ix(\omega-\alpha)}dx = \delta(\omega - \alpha)$$

so we get

$$\int f(x)\overline{g(x)}dx = (\frac{1}{2\pi})^2 \int F(\omega)[\int \overline{G(\alpha)}\delta(\omega - \alpha)d\alpha/2\pi]d\omega/2\pi$$

$$= \int F(\omega)\overline{G(\omega)}d\omega/2\pi.$$

Exercise 10.11: Show that, if f is causal, then R and X are related; specifically, show that X is the *Hilbert transform* of R, that is,

$$X(\omega) = 2\int_{-\infty}^{\infty} \frac{R(\alpha)}{\omega - \alpha}d\alpha.$$

Solution: Since $f(x) = 0$ for $x < 0$ we have $f(x)\text{sgn}(x) = f(x)$. Taking the FT of both sides and applying the convolution theorem, we get

$$F(\omega) = 2i\int F(\alpha)\frac{1}{\omega - \alpha}d\alpha/2\pi.$$

Now compute the real and imaginary parts of both sides.

Exercise 10.12: Compute $\mathcal{F}(z)$ for $f(x) = u(x)$, the Heaviside function. Compare $\mathcal{F}(-i\omega)$ with the FT of u.

Solution: Let $z = a + bi$, where $a > 0$. For $f(x) = u(x)$ the integral becomes

$$\mathcal{F}(z) = \int_0^{\infty} e^{-zx}dx = \frac{-1}{z}[0 - 1] = \frac{1}{z}.$$

Inserting $z = -i\omega$ we get

$$\frac{i}{\omega} = \mathcal{F}(-i\omega) = \int u(x)e^{ix\omega}dx.$$

The integral is the Fourier transform of the Heaviside function $u(x)$, which is not quite equal to $\frac{1}{\omega}$. The point here is that we erroneously evaluated the Laplace transform integral at a point z whose real part is not positive.

Exercise 10.13: Show that if f is radial then its FT F is also radial. Find the FT of the radial function $f(x, y) = \frac{1}{\sqrt{x^2+y^2}}$.

Solution: Inserting $f(r, \theta) = g(r)$ in the equation for $F(\rho, \omega)$ we obtain

$$F(\rho, \omega) = \int_0^\infty \int_{-\pi}^{\pi} g(r)e^{ir\rho\cos(\theta-\omega)}rdrd\theta$$

or

$$F(\rho, \omega) = \int_0^\infty rg(r)[\int_{-\pi}^{\pi} e^{ir\rho\cos(\theta-\omega)}d\theta]dr.$$

Although it does not appear to be, the inner integral is independent of ω; if we replace the variable $\theta - \omega$ with θ we have $\cos\theta$ is the exponent, $d(\theta - \omega) = d\theta$ remains unchanged, and the limits of integration become $-\pi + \omega$ to $\pi + \omega$. But since the integrand is 2π-periodic, this integral is the same as the one from $-\pi$ to π.

To find the FT of the radial function $f(x, y) = \frac{1}{\sqrt{x^2+y^2}}$, we write it in polar coordinates as $f(r, \theta) = g(r) = 1/r$. Then

$$H(\rho) = 2\pi \int_0^\infty J_0(r\rho)dr = \frac{2\pi}{\rho} \int_0^\infty J_0(r\rho)\rho dr = \frac{2\pi}{\rho},$$

since $\int J_0(x)dx = 1$; the basic facts about the Bessel function $J_0(x)$ can be found in most texts on differential equations. So, for the two-dimensional case, the radial function $f(r, \theta) = g(r) = \frac{1}{r}$ is, except for a scaling, its own Fourier transform, as is the case for the standard Gaussian function in one dimension.

Chapter 11: The Uncertainty Principle

Exercise 11.1: Show that, if the inequality is an equation for some f, then $f'(x) = kxf(x)$, so that $f(x) = e^{-\alpha^2 x^2}$ for some $\alpha > 0$.

Solution: We get equality in the Cauchy-Schwarz inequality if and only if

$$f'(x) = cxf(x),$$

for some constant. Solving this differential equation by separation of variables we obtain the solution

$$f(x) = K \exp(\frac{c}{2}x^2).$$

Since we want $\int f(x)dx$ to be finite, we must select $c < 0$.

Chapter 18: Wavelets

Exercise 18.1: Let $u(x) = 1$ for $0 \leq x < \frac{1}{2}$, $u(x) = -1$ for $\frac{1}{2} \leq x < 1$ and zero otherwise. Show that the functions $u_{jk}(x) = u(2^j x - k)$ are mutually orthogonal on the interval $[0, 1]$, where $j = 0, 1, ...$ and $k = 0, 1, ..., 2^j - 1$.

Solution: Consider u_{jk} and u_{mn}, where $m \geq j$. If $m = j$ and $k \neq n$ then the supports are disjoint and the functions are orthogonal. If $m > j$ and the supports are disjoint, then, again, the functions are orthogonal. So suppose that $m > j$ and the supports are not disjoint. Then the support of u_{mn} is a subset of the support of u_{jk}. On that subset $u_{jk}(x)$ is constant, while $u_{mn}(x)$ is that constant for half of the x and is the negative of that constant for the other half; therefore the inner product is zero.

Chapter 21: Fourier Transform Estimation

Exercise 21.1: Use the orthogonality principle to show that the DFT minimizes the distance

$$\int_{-\pi}^{\pi} |F(\omega) - \sum_{m=1}^{M} a_m e^{im\omega}|^2 d\omega.$$

Solution: The orthogonality principle asserts that, for the optimal choice of the a_n, we have

$$\int_{-\pi}^{\pi} (F(\omega) - \sum_{m=1}^{M} a_m e^{im\omega}) e^{-in\omega} d\omega = 0,$$

for $n = 1, ..., M$. It follows, much as in the previous exercise, that $a_n = f(n)$.

Exercise 21.2: Suppose that $0 < \Omega$ and $F(\omega) = 0$ for $|\omega| > \Omega$. Let $f(x)$ be the inverse Fourier transform of $F(\omega)$ and suppose that the data is $f(x_m)$, $m = 1, ..., M$. Use the orthogonality principle to find the coefficients a_m that minimize the distance

$$\int_{-\Omega}^{\Omega} |F(\omega) - \sum_{m=1}^{M} a_m e^{ix_m\omega}|^2 d\omega.$$

Show that the resulting estimate of $F(\omega)$ is consistent with the data.

Solution: The orthogonality principle tells us that, for the optimal choice of the a_m, we have

$$\int_{-\Omega}^{\Omega}(F(\omega-\sum\nolimits_{m=1}^{M}a_m e^{ix_m\omega})e^{-ix_n\omega}d\omega=0,$$

for $n=1,2,...,M$. This says that, for these n,

$$f(x_n)=\sum\nolimits_{m=1}^{M}a_m\int_{-\Omega}^{\Omega}e^{i(x_m-x_n)\omega}d\omega/2\pi$$

or

$$f(x_n)=\sum\nolimits_{m=1}^{M}a_m\frac{\sin\Omega(x_m-x_n)}{\pi(x_m-x_n)}.$$

The inverse Fourier transform of the function

$$F_\Omega(\omega)=\chi_\Omega(\omega)\sum\nolimits_{m=1}^{M}a_m e^{ix_m\omega}$$

is

$$f_\Omega(x)=\sum\nolimits_{m=1}^{M}a_m\frac{\sin\Omega(x_m-x)}{\pi(x_m-x)};$$

setting $x=x_n$ we see that $f_\Omega(x_n)=f(x_n)$, for $n=1,...,M$, so the optimal estimate is data consistent.

Chapter 22: More on Bandlimited Extrapolation

Exercise 22.1: The purpose of this exercise is to show that, for an Hermitian nonnegative-definite M by M matrix Q, a norm-one eigenvector $\mathbf{u}^1$ of Q associated with its largest eigenvalue, λ_1, maximizes the quadratic form $\mathbf{a}^\dagger Q\mathbf{a}$ over all vectors $\mathbf{a}$ with norm one. Let $Q=ULU^\dagger$ be the eigenvector decomposition of Q, where the columns of U are mutually orthogonal eigenvectors $\mathbf{u}^n$ with norms equal to one, so that $U^\dagger U=I$, and $L=diag\{\lambda_1,...,\lambda_M\}$ is the diagonal matrix with the eigenvalues of Q as its entries along the main diagonal. Assume that $\lambda_1\geq\lambda_2\geq...\geq\lambda_M$. Then maximize

$$\mathbf{a}^\dagger Q\mathbf{a}=\sum_{n=1}^{M}\lambda_n\,|\mathbf{a}^\dagger\mathbf{u}^n|^2,$$

subject to the constraint

$$\mathbf{a}^\dagger\mathbf{a}=\mathbf{a}^\dagger U^\dagger U\mathbf{a}=\sum_{n=1}^{M}|\mathbf{a}^\dagger\mathbf{u}^n|^2=1.$$

Solution: Since we have

$$\sum_{n=1}^{M} |\mathbf{a}^{\dagger}\mathbf{u}^{n}|^2 = 1$$

the sum

$$\sum_{n=1}^{M} \lambda_n \, |\mathbf{a}^{\dagger}\mathbf{u}^{n}|^2$$

is a convex combination of the nonnegative numbers λ_n. Such a convex combination must be no greater than the greatest λ_n, which is λ_1. But it can equal λ_1 if we select the unit vector $\mathbf{a}$ to be $\mathbf{a} = \mathbf{u}^1$. So the greatest value $\mathbf{a}^{\dagger}Q\mathbf{a}$ can attain is λ_1.

Exercise 22.2: Show that for the sinc matrix Q_{Ω} the quadratic form $\mathbf{a}^{\dagger}Q\mathbf{a}$ in the previous exercise becomes

$$\mathbf{a}^{\dagger}Q_{\Omega}\mathbf{a} = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} |\sum_{n=1}^{M} a_n e^{in\omega}|^2 d\omega.$$

Show that the norm of the vector $\mathbf{a}$ is the integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |\sum_{n=1}^{M} a_n e^{in\omega}|^2 d\omega.$$

Solution: Write

$$|\sum_{n=1}^{M} a_n e^{in\omega}|^2 = \sum_{n=1}^{M}\sum_{m=1}^{M} a_n \overline{a_m} e^{i(n-m)\omega}.$$

Exercise 22.3: For $M = 30$ compute the eigenvalues of the matrix Q_{Ω} for various choices of Ω, such as $\Omega = \frac{\pi}{k}$, for $k = 2, 3, ..., 10$. For each k arrange the set of eigenvalues in decreasing order and note the proportion of them that are not near zero. The set of eigenvalues of a matrix is sometimes called its *eigenspectrum* and the nonnegative function $\chi_{\Omega}(\omega)$ is a power spectrum; here is one time in which different notions of a *spectrum* are related.

Solution: We find that the eigenvalues separate, more or less, into two groups: those near one and those near zero. The number of eigenvalues in the first group is roughly $30\Omega/\pi$.

Exercise 22.5: Show that the MDFT estimator given by Equation (21.7) can be written as

$$F_{\Omega}(\omega) = \chi_{\Omega}(\omega)\sum_{m=1}^{M} \frac{1}{\lambda_m}(\mathbf{u}^m)^{\dagger}\mathbf{d}\, U^m(\omega),$$

where $\mathbf{d}$ is the data vector.

Solution: Expand $Q^{-1}f$ using the eigenvector/eigenvalue expression for Q^{-1}.

Exercise 22.6: Show that the DFT estimate of $F(\omega)$, restricted to the interval $[-\Omega, \Omega]$, is

$$F_{DFT}(\omega) = \chi_\Omega(\omega) \sum\nolimits_{m=1}^{M} (\mathbf{u}^m)^\dagger \mathbf{d}\, U^m(\omega).$$

Solution: Use the fact that the identity matrix can be written as $I = UU^\dagger$.

Chapter 23: The PDFT

Exercise 23.1: Show that the c_m must satisfy the equations

$$f(x_n) = \sum\nolimits_{m=1}^{M} c_m p(x_n - x_m),\ n = 1, ..., M,$$

where $p(x)$ is the inverse Fourier transform of $P(\omega)$.

Solution: The inverse FT of the function $F_{PDFT}(\omega)$ is

$$f_{PDFT}(x) = \sum\nolimits_{m=1}^{M} c_m p(x - x_m).$$

In order for $f_{PDFT}(x)$ to be data consistent we must have

$$f_{PDFT}(x_n) = \sum\nolimits_{m=1}^{M} c_m p(x_n - x_m)$$

for $n = 1, ..., M$.

Exercise 23.2: Show that the estimate $F_{PDFT}(\omega)$ minimizes the distance

$$\int |F(\omega) - P(\omega) \sum\nolimits_{m=1}^{M} a_m \exp(ix_m\omega)|^2 P(\omega)^{-1} d\omega$$

over all choices of the coefficients a_m.

Solution: According to the orthogonality principle the optimal choice $a_m = c_m$ must satisfy

$$0 = \int (F(\omega) - P(\omega) \sum\nolimits_{m=1}^{M} c_m \exp(ix_m\omega)) P(\omega) e^{-ix_n\omega} P(\omega)^{-1} d\omega,$$

for $n = 1, ..., M$. Therefore

$$0 = \int (F(\omega) - P(\omega) \sum_{m=1}^{M} c_m \exp(ix_m\omega))e^{-ix_n\omega} d\omega,$$

which tells us that

$$f(x_n) = \sum_{m=1}^{M} c_m p(x_n - x_m)$$

for $n = 1, ..., M$.

Chapter 24: A Little Matrix Theory

Exercise 24.1: Show that if $\mathbf{z} = (z_1, ..., z_N)^T$ is a column vector with complex entries and $H = H^\dagger$ is an N by N Hermitian matrix with complex entries then the quadratic form $\mathbf{z}^\dagger H\mathbf{z}$ is a real number. Show that the quadratic form $\mathbf{z}^\dagger H\mathbf{z}$ can be calculated using only real numbers. Let $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, with $\mathbf{x}$ and $\mathbf{y}$ real vectors and let $H = A + iB$, where A and B are real matrices. Then show that $A^T = A$, $B^T = -B$, $\mathbf{x}^T B\mathbf{x} = 0$ and finally,

$$\mathbf{z}^\dagger H\mathbf{z} = [\,\mathbf{x}^T \quad \mathbf{y}^T\,] \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}.$$

Use the fact that $\mathbf{z}^\dagger H\mathbf{z}$ is real for every vector $\mathbf{z}$ to conclude that the eigenvalues of H are real.

Solution: The quadratic form $\mathbf{z}^\dagger H\mathbf{z}$ is a complex number and also the product of three matrices. Its conjugate transpose is simply its complex conjugate, since it is only 1 by 1; but

$$(\mathbf{z}^\dagger H\mathbf{z})^\dagger = \mathbf{z}^\dagger H^\dagger (\mathbf{z}^\dagger)^\dagger = \mathbf{z}^\dagger H\mathbf{z}$$

since H is Hermitian. The complex conjugate of $\mathbf{z}^\dagger H\mathbf{z}$ is itself, so it must be real. We have

$$A + iB = H = H^\dagger = A^T - iB^T,$$

so that $A = A^T$ and $B^T = -B$.

Writing $\mathbf{z}^\dagger Q\mathbf{z}$ in terms of A, B, $\mathbf{x}$ and $\mathbf{y}$ we get

$$\mathbf{z}^\dagger Q\mathbf{z} = (\mathbf{x}^T - i\mathbf{y}^T)(A + iB)(\mathbf{x} + i\mathbf{y}) = (\mathbf{x}^T - i\mathbf{y}^T)(A\mathbf{x} - B\mathbf{y} + i(B\mathbf{x} + A\mathbf{y})$$

$$= \mathbf{x}^T A\mathbf{x} - \mathbf{x}^T B\mathbf{y} + \mathbf{y}^T B\mathbf{x} + \mathbf{y}^T A\mathbf{y} + i(\mathbf{x}^T B\mathbf{x} + \mathbf{x}^T A\mathbf{y} - \mathbf{y}^T A\mathbf{x} + \mathbf{y}^T B\mathbf{y})$$

$$= \mathbf{x}^T A\mathbf{x} + \mathbf{y}^T A\mathbf{y} - \mathbf{x}^T B\mathbf{y} + \mathbf{y}^T B\mathbf{x}$$

since

$$\mathbf{x}^T B\mathbf{x} = (\mathbf{x}^T B\mathbf{x})^T = \mathbf{x}^T B^T\mathbf{x} = -\mathbf{x}^T B\mathbf{x}$$

implies that $\mathbf{x}^T B\mathbf{x} = 0$ and, similarly, $\mathbf{y}^T B\mathbf{y} = 0$.

Let λ be an eigenvalue of H associated with eigenvector $\mathbf{u}$. Then

$$\mathbf{u}^\dagger H\mathbf{u} = \mathbf{u}^\dagger(\lambda\mathbf{u}) = \lambda\mathbf{u}^\dagger\mathbf{u} = \lambda.$$

Since $\mathbf{u}^\dagger H\mathbf{u}$ is real, so is λ.

Exercise 24.2: Let A be an M by N matrix with complex entries. View A as a linear function with domain C^N, the space of all N-dimensional complex column vectors, and range contained within C^M, via the expression $A(\mathbf{x}) = A\mathbf{x}$. Suppose that $M > N$. The range of A, denoted $R(A)$, cannot be all of C^M. Show that every vector $\mathbf{z}$ in C^M can be written uniquely in the form $\mathbf{z} = A\mathbf{x} + \mathbf{w}$, where $A^\dagger\mathbf{w} = \mathbf{0}$. Show that $\|\mathbf{z}\|^2 = \|A\mathbf{x}\|^2 + \|\mathbf{w}\|^2$, where $\|\mathbf{z}\|^2$ denotes the square of the norm of $\mathbf{z}$. Hint: If $\mathbf{z} = A\mathbf{x} + \mathbf{w}$ then consider $A^\dagger\mathbf{z}$. Assume $A^\dagger A$ is invertible.

Solution: We assume that $A^\dagger A$ is invertible. If $\mathbf{z} = A\mathbf{x} + \mathbf{v}$ with $A^\dagger\mathbf{v} = 0$ then $A^\dagger\mathbf{z} = A^\dagger A\mathbf{x}$, so that $\mathbf{x} = (A^\dagger A)^{-1}A^\dagger\mathbf{z}$. Then

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

and we see easily that $A^\dagger\mathbf{v} = 0$. Then we have

$$||\mathbf{z}||^2 = ||A\mathbf{x} + \mathbf{v}||^2 = \mathbf{x}^\dagger A^\dagger A\mathbf{x} + \mathbf{x}^\dagger A^\dagger\mathbf{v} + \mathbf{v}^\dagger A\mathbf{x} + \mathbf{v}^\dagger\mathbf{v} = ||A\mathbf{x}||^2 + ||\mathbf{v}||^2.$$

Exercise 24.5: Show that the nonzero eigenvalues of A and B are the same.

Solution: Let λ be a nonzero eigenvalue of A, with $A\mathbf{u} = \lambda\mathbf{u}$ for some nonzero vector $\mathbf{u}$. Then $CA\mathbf{u} = \lambda C\mathbf{u}$ or $(CC^\dagger)C\mathbf{u} = BC\mathbf{u} = \lambda C\mathbf{u}$; with $C\mathbf{u} = \mathbf{v}$ we have $B\mathbf{v} = \lambda\mathbf{v}$. Since B is invertible $\mathbf{v}$ is not the zero vector. So λ is an eigenvalue of B.

Conversely, let $\lambda \neq 0$ be an eigenvalue of B, with $B\mathbf{v} = \lambda\mathbf{v}$ for some nonzero $\mathbf{v}$. Then $B\mathbf{v} = CC^\dagger\mathbf{v} = \lambda\mathbf{v}$ and so $C^\dagger B\mathbf{v} = (C^\dagger C)C^\dagger\mathbf{v} = AC^\dagger\mathbf{v} = \lambda C^\dagger\mathbf{v}$. We need to show that $\mathbf{w} = C^\dagger\mathbf{v}$ is not the zero vector. If $\mathbf{0} = \mathbf{w} = C^\dagger\mathbf{v}$ then $\mathbf{0} = C\mathbf{w} = CC^\dagger\mathbf{v} = B\mathbf{v}$. But B is invertible and $\mathbf{v}$ is nonzero; this is a contradiction, so we conclude that $\mathbf{w} \neq \mathbf{0}$.

Exercise 24.6: Show that $UMV^\dagger$ equals C.

Solution: The first N columns of the matrix UM form the matrix

$$ULL^{-1/2} = BUL^{-1/2}$$

and the remaining columns are zero. Consider the product $V(UM)^\dagger$. The first N columns of V form the matrix $C^\dagger UL^{-1/2}$ so

$$V(UM)^\dagger = C^\dagger UL^{-1}U^\dagger B = C^\dagger B^{-1}B = C^\dagger$$

and so $UMV^\dagger = C$.

Exercise 24.7: If $N > K$ the system $C\mathbf{x} = \mathbf{d}$ probably has no exact solution. Show that $C^* = (C^\dagger C)^{-1}C^\dagger$ so that the vector $\mathbf{x} = C^*\mathbf{d}$ is the least squares approximate solution.

Solution: Show that $(C^\dagger C)C^* = C^\dagger = VM^TU^\dagger$.

Exercise 24.8: If $N < K$ the system $C\mathbf{x} = \mathbf{d}$ probably has infinitely many solutions. Show that the pseudo-inverse is now $C^* = C^\dagger(CC^\dagger)^{-1}$, so that the vector $\mathbf{x} = C^*\mathbf{d}$ is the exact solution of $C\mathbf{x} = \mathbf{d}$ closest to the origin; that is, it is the minimum norm solution.

Solution: Show that $C^*(CC^\dagger) = C^\dagger$.

Exercise 24.9: Show that the vector $\mathbf{x} = (x_1, ..., x_N)^T$ minimizes the mean squared error

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \sum_{m=1}^{N}(A\mathbf{x}_m - b_m)^2,$$

if and only if $\mathbf{x}$ satisfies the system of linear equations $A^T(A\mathbf{x} - \mathbf{b}) = \mathbf{0}$, where $A\mathbf{x}_m = (A\mathbf{x})_m = \sum_{n=1}^{N} A_{mn}x_n$. Hint: calculate the partial derivative of $\|A\mathbf{x} - \mathbf{b}\|^2$ with respect to each x_n.

Solution: The partial derivative of $\|A\mathbf{x} - \mathbf{b}\|^2$ with respect to x_n is

$$2\sum\nolimits_{m=1}^{M} A_{mn}(A\mathbf{x}_m - b_m).$$

Setting each of these partial derivatives equal to zero gives

$$A^T(A\mathbf{x} - \mathbf{b}) = 0.$$

Exercise 24.11: Show that any vector $\mathbf{p}$ can be written as $\mathbf{p} = A^T\mathbf{q} + \mathbf{r}$, where $A\mathbf{r} = 0$.

Solution: If $M \leq N$, then, by our assumption, AA^T is invertible. If the decomposition of $\mathbf{p}$ does hold,, we can calculate the $\mathbf{q}$ and $\mathbf{r}$ as follows. Multiply both sides of the equation by A, to obtain $A\mathbf{p} = AA^T\mathbf{q} + A\mathbf{r} =$

$AA^T\mathbf{q}$. Solving for $\mathbf{q}$, we have $\mathbf{q} = (AA^T)^{-1}A\mathbf{p}$. It follows that $\mathbf{r} = \mathbf{p} - A^T(AA^T)^{-1}A\mathbf{p}$, which clearly has the property $A\mathbf{r} = 0$. Now without assuming the truth of the statement, we can still write any $\mathbf{p}$ as

$$\mathbf{p} = A^T(AA^T)^{-1}A\mathbf{p} + \mathbf{p} - A^T(AA^T)^{-1}A\mathbf{p}.$$

If $M > N$, then, since we are assuming that A has full rank, the range of A^T must be all of R^N, so every $\mathbf{p}$ in R^N has the form $\mathbf{p} = A^T\mathbf{q}$, for some $\mathbf{q}$.

Exercise 24.12: Show that F_ϵ always has a unique minimizer $\hat{\mathbf{x}}_\epsilon$ given by

$$\hat{\mathbf{x}}_\epsilon = ((1-\epsilon)A^TA + \epsilon I)^{-1}((1-\epsilon)A^T\mathbf{b} + \epsilon\mathbf{p});$$

this is a regularized solution of $A\mathbf{x} = \mathbf{b}$. Here $\mathbf{p}$ is a prior estimate of the desired solution. Note that the inverse above always exists.

Solution: Set to zero the partial derivatives with respect to each of the variables x_n. Show that the second derivative matrix is $A^TA + \epsilon I$, which is positive-definite; therefore the partial derivatives are zero at a minimum.

Exercise 24.13: Show that, in **Case 1**, taking limits as $\epsilon \to 0$ on both sides of the expression for $\hat{\mathbf{x}}_\epsilon$ gives $\hat{\mathbf{x}}_\epsilon \to (A^TA)^{-1}A^T\mathbf{b}$, the least squares solution of $A\mathbf{x} = \mathbf{b}$.

Solution: In this case we can simply set $\epsilon = 0$, since the inverse $(A^TA)^{-1}$ exists.

Exercise 24.14: Show that

$$((1-\epsilon)A^TA + \epsilon I)^{-1}(\epsilon\mathbf{r}) = \mathbf{r}, \forall\epsilon.$$

Solution: As in the hint, let

$$\mathbf{t}_\epsilon = ((1-\epsilon)A^TA + \epsilon I)^{-1}(\epsilon\mathbf{r}).$$

Then multiplying by A gives

$$A\mathbf{t}_\epsilon = A((1-\epsilon)A^TA + \epsilon I)^{-1}(\epsilon\mathbf{r}).$$

Now it follows from $A\mathbf{r} = \mathbf{0}$ and

$$((1-\epsilon)AA^T + \epsilon I)^{-1}A = A((1-\epsilon)A^TA + \epsilon I)^{-1}$$

that $A\mathbf{t}_\epsilon = \mathbf{0}$. Now multiply both sides of the equation

$$\mathbf{t}_\epsilon = ((1-\epsilon)A^TA + \epsilon I)^{-1}(\epsilon\mathbf{r})$$

by $(1-\epsilon)A^TA+\epsilon I$ to get $\epsilon\mathbf{t}_\epsilon = \epsilon\mathbf{r}$. Now we take the limit of $\hat{\mathbf{x}}_\epsilon$, as $\epsilon \to 0$, by setting $\epsilon = 0$, to get $\hat{\mathbf{x}}_\epsilon \to A^T(AA^T)^{-1}\mathbf{b}+\mathbf{r} = \hat{\mathbf{x}}$.

Now we show that $\hat{\mathbf{x}}$ is the solution of $A\mathbf{x} = \mathbf{b}$ closest to $\mathbf{p}$. By the orthogonality theorem it must then be the case that $\langle \mathbf{p}-\hat{\mathbf{x}}, \mathbf{x}-\hat{\mathbf{x}}\rangle = \mathbf{0}$ for every $\mathbf{x}$ with $A\mathbf{x} = \mathbf{b}$. Since $\mathbf{p}-\hat{\mathbf{x}} = A^T\mathbf{q} - A^T(AA^T)^{-1}\mathbf{b}$ we have

$$\langle \mathbf{p}-\hat{\mathbf{x}}, \mathbf{x}-\hat{\mathbf{x}}\rangle = \langle \mathbf{q}-(AA^T)^{-1}\mathbf{b}, A\mathbf{x}-A\hat{\mathbf{x}}\rangle = 0.$$

Chapter 25: Matrix and Vector Differentiation

Exercise 25.1: Let $\mathbf{y}$ be a fixed real column vector and $z = f(\mathbf{x}) = \mathbf{y}^T\mathbf{x}$. Show that

$$\frac{\partial z}{\partial \mathbf{x}} = \mathbf{y}.$$

Solution: We write

$$z = \mathbf{y}^T\mathbf{x} = \sum_{n=1}^N x_n y_n$$

so that

$$\frac{\partial z}{\partial x_n} = y_n$$

for each n.

Exercise 25.2: Let Q be a real symmetric nonnegative definite matrix and let $z = f(\mathbf{x}) = \mathbf{x}^TQ\mathbf{x}$. Show that the gradient of this quadratic form is

$$\frac{\partial z}{\partial \mathbf{x}} = 2Q\mathbf{x}.$$

Solution: Following the hint, we write Q as a linear combination of dyads involving the eigenvectors; that is

$$Q = \sum_{m=1}^N \lambda_m \mathbf{u}^m(\mathbf{u}^m)^\dagger.$$

Then

$$z = \mathbf{x}^TQ\mathbf{x} = \sum_{m=1}^N \lambda_m(\mathbf{x}^T\mathbf{u}^m)^2$$

so that

$$z = \sum_{m=1}^N \lambda_m(\sum_{n=1}^N x_n u_n^m)^2.$$

Therefore, the partial derivative of z with respect to x_n is

$$\frac{\partial z}{\partial x_n} = 2\sum_{m=1}^N \lambda_n(x_n u_m^n)u_m^n,$$

which can then be written as

$$\frac{\partial z}{\partial \mathbf{x}} = 2Q\mathbf{x}.$$

Exercise 25.3: Let $z = ||A\mathbf{x} - \mathbf{b}||^2$. Show that

$$\frac{\partial z}{\partial \mathbf{x}} = 2A^T A\mathbf{x} - 2A^T\mathbf{b}.$$

Solution: Using $z = (A\mathbf{x} - \mathbf{b})^T(A\mathbf{x} - \mathbf{b})$ we get

$$z = \mathbf{x}^T A^T A\mathbf{x} - \mathbf{b}^T A\mathbf{x} - \mathbf{x}^T A^T\mathbf{b} + \mathbf{b}^T\mathbf{b}.$$

Then it follows from the two previous exercises that

$$\frac{\partial z}{\partial \mathbf{x}} = 2A^T A\mathbf{x} - 2A^T\mathbf{b}.$$

Exercise 25.4: Suppose $(u, v) = (u(x, y), v(x, y))$ is a change of variables from the Cartesian (x, y) coordinate system to some other (u, v) coordinate system. Let $\mathbf{x} = (x, y)^T$ and $\mathbf{z} = (u(\mathbf{x}), v(\mathbf{x}))^T$.

a: Calculate the Jacobian for the rectangular coordinate system obtained by rotating the (x, y) system through an angle of θ.

Solution: The equations for this change of coordinates are

$$u = x\cos\theta + y\sin\theta,$$

and

$$v = -x\sin\theta + y\cos\theta.$$

Then $u_x = \cos\theta$, $u_y = \sin\theta$, $v_x = -\sin\theta$ and $v_y = \cos\theta$. The Jacobian is therefore one.

b: Calculate the Jacobian for the transformation from the (x, y) system to polar coordinates.

Solution: We have $r = \sqrt{x^2 + y^2}$ and $\tan\theta = \frac{y}{x}$. Writing $r^2 = x^2 + y^2$, we get $2rr_x = 2x$ and $2rr_y = 2y$, so that $r_x = x/r$ and $r_y = y/r$. Also

$$(\sec\theta)^2\theta_x = -y/x^2$$

and

$$(\sec\theta)^2\theta_y = 1/x.$$

Since $\sec\theta = r/x$ we get

$$\theta_x = \frac{x^2}{r^2}\frac{-y}{x^2} = \frac{-y}{r^2}$$

and

$$\theta_y = \frac{x^2}{r^2}\frac{1}{x} = \frac{x}{r^2}.$$

The Jacobian is therefore $\frac{1}{r}$.

Exercise 25.6: Show that the derivative of $z = \mathrm{trace}\,(DAC)$ with respect to A is

$$\frac{\partial z}{\partial A} = D^T C^T.$$

Solution: Just write out the general term of DAC.

Exercise 25.7: Let $z = \mathrm{trace}\,(A^T CA)$. Show that the derivative of z with respect to the matrix A is

$$\frac{\partial z}{\partial A} = CA + C^T A.$$

Therefore, if $C = Q$ is symmetric, then the derivative is $2QA$.

Solution: Again, just write out the general term of $A^T CA$.

Chapter 27: Discrete Random Processes

Exercise 27.1: Show that the autocorrelation matrix R is nonnegative definite. Under what conditions can R fail to be positive-definite?

Solution: Let

$$A(\omega) = \sum\nolimits_{n=1}^{N+1} a_n e^{in\omega}.$$

Then we have

$$\int |A(\omega)|^2 R(\omega) d\omega = \mathbf{a}^\dagger R\mathbf{a} \geq 0.$$

If the quadratic form $\mathbf{a}^\dagger R\mathbf{a} = 0$ for some vector $\mathbf{a}$ then the integral must also be zero, which says that the power spectrum is nonzero only when the polynomial is zero; that is, the power spectrum $R(\omega)$ is a sum of not more than N delta functions.

Chapter 28: Best Linear Unbiased Estimation

Exercise 28.1: Show that

$$E(|\hat{\mathbf{x}} - \mathbf{x}|^2) = \text{trace} K^\dagger Q K.$$

Solution: Write the left side as

$$E(\text{trace}\,((\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^\dagger)).$$

Also use the fact that the trace and expected value operations commute. Then

$$E(|\hat{\mathbf{x}} - \mathbf{x}|^2) = \text{trace}(E(K^\dagger \mathbf{z}\mathbf{z}^\dagger K - \mathbf{x}\mathbf{z}^\dagger K - K^\dagger \mathbf{z}\mathbf{x}^\dagger + \mathbf{x}\mathbf{x}^\dagger)) = E(K^\dagger \mathbf{z}\mathbf{z}^\dagger K) - \mathbf{x}\mathbf{x}^\dagger.$$

Notice that

$$\mathbf{z}\mathbf{z}^\dagger = H\mathbf{x}\mathbf{x}^\dagger H^\dagger + H\mathbf{x}\mathbf{v}^\dagger + \mathbf{v}\mathbf{x}^\dagger H^\dagger + \mathbf{v}\mathbf{v}^\dagger.$$

Therefore

$$E(K^\dagger \mathbf{z}\mathbf{z}^\dagger K) = K^\dagger H\mathbf{x}\mathbf{x}^\dagger H^\dagger K + K^\dagger Q K.$$

It follows that

$$E(|\hat{\mathbf{x}} - \mathbf{x}|^2) = \text{trace} K^\dagger Q K.$$

Chapter 29: The BLUE and the Least Squares Estimators:

Exercise 29.4: Let Q be Hermitian. Show that $Q(S)^\perp = Q^{-1}(S^\perp)$ for every subspace S. If Q is also invertible then $Q^{-1}(S)^\perp = Q(S^\perp)$. Find an example of a non-invertible Q for which $Q^{-1}(S)^\perp$ and $Q(S^\perp)$ are different.

Solution: First, we show that $Q(S)^\perp = Q^{-1}(S^\perp)$. Suppose that $a \in Q(S)^\perp$. Then, $a^\dagger Q s = 0$, for all $s \in S$, so that $a^\dagger Q^\dagger s = (Qa)^\dagger s = 0$, for all $s \in S$. Therefore, $Qa \in S^\perp$ and $a \in Q^{-1}(S^\perp)$. Conversely, if $b \in Q^{-1}(S^\perp)$, then $Qb \in S^\perp$, or $b \in Q^{-1}(S^\perp)$.

Now suppose that Q is invertible. Then we can use Q^{-1} in place of Q in the first part of this exercise. The desired result follows immediately. If Q is not invertible, however, the result may not hold, as the following example proves.

Let Q be the matrix that transforms any vector $(a, b, c)^T \in R^3$ to the vector $(0, a, b)^T$; this Q is not invertible. Let S be the subspace consisting of all vectors of the form $(0, b, c)^T$. Then, $Q^{-1}(S) = R^3$, so that $Q^{-1}(S)^\perp = \{0\}$. But, $Q(S^\perp)$ is the subspace of vectors of the form $(0, a, 0)^T$.

Chapter 31: The Vector Wiener Filter

Exercise 31.1: Apply the vector Wiener filter to the simplest problem discussed earlier. Here let $K = 1$ and $NN^\dagger = Q$.

Solution: Let $\mathbf{1} = (1, 1, ..., 1)^T$, so that the signal vector is $\mathbf{s} = c\mathbf{1}$ for some constant c and the data vector is $\mathbf{z} = c\mathbf{1} + \mathbf{v}$. Then $SS^\dagger = \mathbf{1}\mathbf{1}^T$. We have

$$(Q + \mathbf{1}\mathbf{1}^\dagger)^{-1} = Q^{-1} - (1 + \mathbf{1}^\dagger Q^{-1}\mathbf{1})^{-1} Q^{-1}\mathbf{1}\mathbf{1}^\dagger Q^{-1},$$

so we get

$$\hat{\mathbf{s}} = \frac{\mathbf{1}^\dagger Q^{-1}\mathbf{z}}{1 + \mathbf{1}^\dagger Q^{-1}\mathbf{1}}\mathbf{1},$$

and the estimate of the constant c is

$$\hat{c} = \frac{\mathbf{1}^\dagger Q^{-1}\mathbf{z}}{1 + \mathbf{1}^\dagger Q^{-1}\mathbf{1}}.$$

When the noise power is very low the denominator is dominated by the second term and we get the BLUE estimate.

Exercise 31.3 Assume that $E(\mathbf{x}\mathbf{x}^\dagger) = \sigma^2 I$. Show that the mean squared error for the VWF estimate is

$$E(|\hat{\mathbf{s}} - \mathbf{s}|^2) = \text{trace}\,(H(H^\dagger Q^{-1}H + \sigma^{-2}I)^{-1}H^\dagger).$$

Solution: We know from the calculations on p. 182 that, for any B,

$$E|\hat{\mathbf{s}} - \mathbf{s}|^2 = \text{trace}\,(B^\dagger(R_s + R_v)B - R_sB - B^\dagger R_s + R_s).$$

The optimal B is $B = (R_s + R_v)^{-1}R_s$, so that, for this B, we have

$$E|\hat{\mathbf{s}} - \mathbf{s}|^2 = \text{trace}\,(R_s - R_s(R_s + R_v)^{-1}R_s).$$

With $E(\mathbf{x}\mathbf{x}^\dagger) = \sigma^2 I$, it follows that $R_s = HH^\dagger$. Let $Q = R_v$. Inserting the optimal

$$B = (\sigma^2 I + Q)^{-1}(\sigma^2 I) = \sigma^2(\sigma^2 I + Q)^{-1},$$

we get

$$E|\hat{\mathbf{s}} - \mathbf{s}|^2 = \text{trace}\,(\sigma^2 HH^\dagger - \sigma^4 HH^\dagger(Q + HH^\dagger)^{-1}HH^\dagger).$$

Now use the identity

$$\sigma^2 H^\dagger(Q + \sigma^2 HH^\dagger)^{-1} = (H^\dagger Q^{-1}H + \sigma^{-2}I)^{-1}H^\dagger Q^{-1},$$

and the fact that

$$\begin{aligned}
&(H^\dagger Q^{-1}H + \sigma^{-2}I)^{-1}H^\dagger Q^{-1}H \\
&= (H^\dagger Q^{-1}H + \sigma^{-2}I)^{-1}(H^\dagger Q^{-1}H + \sigma^{-2}I - \sigma^{-2}I) \\
&= I - \sigma^{-2}(H^\dagger Q^{-1}H + \sigma^{-2}I)^{-1}.
\end{aligned}$$

Chapter 34: Entropy Maximization

Exercise 34.1 What happened in Figure 34.3?

Solution: The points on the horizontal axis that were used to make the graph were not sufficiently dense to capture the very sharp peaks in the actual answer. While this is not a common problem in practice, it can occur in simulation studies, and may puzzle those not accustomed to graphing high-resolution estimates.

Chapter 37: Eigenvector Methods

Exercise 37.2: Show that $\lambda_m = \sigma^2$ for $m = J+1, ..., M$, while $\lambda_m > \sigma^2$ for $m = 1, ..., J$.

Solution: From Exercise 37.1 we conclude that, for any vector $\mathbf{u}$ the quadratic form $\mathbf{u}^\dagger R\mathbf{u}$ is

$$\mathbf{u}^\dagger R\mathbf{u} = \sum_{j=1}^{J} |A_j|^2 |\mathbf{u}^\dagger \mathbf{e}_j|^2 + \sigma^2 |\mathbf{u}^\dagger \mathbf{u}|^2.$$

The norm-one eigenvectors of R associated with the J largest eigenvalues will lie in the linear span of the vectors $\mathbf{e}_j$, $j = 1, ..., J$, while the remaining $M - J$ eigenvectors will be orthogonal to the $\mathbf{e}_j$. For these remaining eigenvectors the quadratic form will have the value $\lambda_m = \sigma^2$, since the eigenvectors have norm equal to one. For the eigenvectors associated with the J largest eigenvalues, the quadratic form will be greater than σ^2, since it will also involve a positive term coming from the sum.

Since $M > J$ the $M - J$ orthogonal eigenvectors $\mathbf{u}_m$ corresponding to λ_m for $m = J+1, ..., M$ will be orthogonal to each of the $\mathbf{e}_j$. Then consider the quadratic forms $\mathbf{u}_m^\dagger R\mathbf{u}_m$.

Chapter 39: Some Probability Theory

Exercise 39.1: Prove these two assertions.

Solution: The expected value of $\overline{X}$ is

$$E(\overline{X}) = \frac{1}{N} \sum_{n=1}^{N} E(X_n) = \frac{1}{N} \sum_{n=1}^{N} \mu = \mu.$$

The variance of $\overline{X}$ is

$$E((\overline{X} - \mu)^2) = E(\overline{X}^2 - 2\mu\overline{X} + \mu^2)$$

$$= E(\overline{X}^2) - \mu^2.$$

Then

$$E(\overline{X}^2) = \frac{1}{N^2} E(\sum_{n=1}^{N} X_n \sum_{m=1}^{N} X_m).$$

Now use the fact that $E(X_n X_m) = E(X_n)E(X_m) = \mu^2$ if $m \neq n$ while $E(X_n X_n) = \sigma^2 + \mu^2$.

Exercise 39.3: Show that the sequence $\{p_k\}_{k=0}^{\infty}$ sums to one.

Solution: The Taylor series expansion of the function e^x is

$$e^x = \sum\nolimits_{k=0}^{\infty} \frac{x^k}{k!},$$

so

$$\sum\nolimits_{k=0}^{\infty} p_k = e^{-\lambda} \sum\nolimits_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1.$$

Exercise 39.4: Show that the expected value $E(X)$ is λ, where the expected value in this case is

$$E(X) = \sum\nolimits_{k=0}^{\infty} k p_k.$$

Solution: Note that

$$\sum_{k=0}^{\infty} k p_k = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}$$

$$= e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda.$$

Exercise 39.5: Show that the variance of X is also λ, where the variance of X in this case is

$$\text{var}(X) = \sum\nolimits_{k=0}^{\infty} (k-\lambda)^2 p_k.$$

Solution: Use

$$(k-\lambda)^2 = k^2 - 2k\lambda + \lambda^2 = k(k-1) + k - 2k\lambda + \lambda^2.$$

Chapter 42: Signal Detection and Estimation

Exercise 42.1: Use Cauchy's inequality to show that, for any fixed vector $\mathbf{a}$, the choice $\mathbf{b} = \beta\mathbf{a}$ maximizes the quantity $|\mathbf{b}^\dagger\mathbf{a}|^2/\mathbf{b}^\dagger\mathbf{b}$, for any constant β.

Solution: According to Cauchy's inequality the quantity $\frac{|\mathbf{b}^\dagger\mathbf{a}|^2}{\mathbf{b}^\dagger\mathbf{b}}$ does not exceed $\mathbf{a}^\dagger\mathbf{a}$. The choice of $\mathbf{b} = \beta\mathbf{a}$ makes the ratio equal to $\mathbf{a}^\dagger\mathbf{a}$, so maximizes the ratio.

Exercise 42.2: Use the definition of the correlation matrix Q to show that Q is Hermitian and that, for any vector $\mathbf{y}$, $\mathbf{y}^\dagger Q\mathbf{y} \geq 0$. Therefore Q is a nonnegative definite matrix and, using its eigenvector decomposition, can be written as $Q = CC^\dagger$, for some invertible square matrix C.

Solution: The entry of Q in the m-th row and n-th column is $Q_{mn} = E(z_m\overline{z_n})$, so $Q_{nm} = \overline{Q_{mn}}$. For any vector $\mathbf{y}$ the quadratic form $\mathbf{y}^\dagger Q\mathbf{y} = E(|\mathbf{y}^\dagger\mathbf{z}|^2)$ and the expected value of a nonnegative random variable is nonnegative. Therefore Q is Hermitian and nonnegative-definite, so its eigenvalues are nonnegative. The eigenvector/eigenvalue decomposition is $Q = ULU^\dagger$, where L is the diagonal matrix with the eigenvalues on the main diagonal. Since these eigenvalues are nonnegative, they have nonnegative square roots. Make these the diagonal elements of the matrix $L^{1/2}$ and write $C = UL^{1/2}U^\dagger$. Then we have $C = C^\dagger$ and $CC^\dagger = C^\dagger C = Q$.

Exercise 42.3: Consider now the problem of maximizing $|\mathbf{b}^\dagger\mathbf{s}|^2/\mathbf{b}^\dagger Q\mathbf{b}$. Using the two previous exercises, show that the solution is $\mathbf{b} = \beta Q^{-1}\mathbf{s}$, for some arbitrary constant β.

Solution: Write $\mathbf{b}^\dagger Q\mathbf{b} = \mathbf{b}^\dagger C^\dagger C\mathbf{b} = \mathbf{d}^\dagger\mathbf{d}$, for $\mathbf{d} = C\mathbf{b}$. We assume that Q is invertible, so C is also. Write

$$\mathbf{b}^\dagger\mathbf{s} = \mathbf{b}^\dagger C^\dagger(C^\dagger)^{-1}\mathbf{s} = \mathbf{d}^\dagger\mathbf{e},$$

for $\mathbf{e} = (C^\dagger)^{-1}\mathbf{s}$. So the problem now is to maximize the ratio $\frac{|\mathbf{d}^\dagger\mathbf{e}|^2}{\mathbf{d}^\dagger\mathbf{d}}$. By the first exercise we know that this ratio is maximized when we select $\mathbf{d} = \beta\mathbf{e}$ for some constant β. This means that $C\mathbf{b} = \beta(C^\dagger)^{-1}\mathbf{s}$ or $\mathbf{b} = \beta Q^{-1}\mathbf{s}$. Here the β is a free choice; we select it so that $\mathbf{b}^\dagger\mathbf{s} = 1$.

Chapter 47: A Tale of Two Algorithms

Exercise 47.1: Show that

$$KL(\mathbf{x},\mathbf{z}) = KL(x_+, z_+) + KL(\mathbf{x}, \frac{x_+}{z_+}\mathbf{z})$$

for any nonnegative vectors $\mathbf{x}$ and $\mathbf{z}$, with x_+ and $z_+ > 0$ denoting the sums of the entries of vectors $\mathbf{x}$ and $\mathbf{z}$, respectively.

Solution: Begin with $KL(\mathbf{x}, \frac{x_+}{z_+}\mathbf{z})$ and write it out as

$$KL(\mathbf{x}, \frac{x_+}{z_+}\mathbf{z}) = \sum_{n=1}^{N} x_n \log(x_n / \frac{x_+}{z_+} z_n) + \frac{x_+}{z_+} \sum_{n=1}^{N} z_n - \sum_{n=1}^{N} x_n$$

$$= \sum_{n=1}^{N} (x_n \log \frac{x_n}{z_n} + z_n - x_n) - \sum_{n=1}^{N} (x_n \log \frac{x_+}{z_+} + (\frac{x_+}{z_+} - 1) z_n)$$

$$= KL(\mathbf{x}, \mathbf{z}) - x_+ \log \frac{x_+}{z_+} + x_+ - z_+ = KL(\mathbf{x}, \mathbf{z}) - KL(x_+, z_+).$$

Chapter 50: The Algebraic Reconstruction Technique

Exercise 50.1: Establish the following facts concerning the ART.

Fact 1:

$$||\mathbf{x}^k||^2 - ||\mathbf{x}^{k+1}||^2 = (A(\mathbf{x}^k)_{m(k)})^2 - (b_{m(k)})^2.$$

Solution: Write $||\mathbf{x}^{k+1}||^2 = ||\mathbf{x}^k + (\mathbf{x}^{k+1} - \mathbf{x}^k)||^2$ and expand using the complex dot product.

Fact 2:

$$||\mathbf{x}^{rM}||^2 - ||\mathbf{x}^{(r+1)M}||^2 = ||\mathbf{v}^r||^2 - ||\mathbf{b}||^2.$$

Solution: The solution is similar to that of the previous exercise.

Fact 3:

$$||\mathbf{x}^k - \mathbf{x}^{k+1}||^2 = ((A\mathbf{x}^k)_{m(k)} - b_m)^2.$$

Solution: Easy.

Fact 4: There exists $B > 0$ such that, for all $r = 0, 1, ...,$ if $||\mathbf{v}^r|| \leq ||\mathbf{b}||$ then $||\mathbf{x}^{rM}|| \geq ||\mathbf{x}^{(r+1)M}|| - B$.

Solution: This is an application of the triangle inequality.

Fact 5: Let $\mathbf{x}^0$ and $\mathbf{y}^0$ be arbitrary and $\{\mathbf{x}^k\}$ and $\{\mathbf{y}^k\}$ the sequences generated by applying the ART algorithm. Then

$$||\mathbf{x}^0 - \mathbf{y}^0||^2 - ||\mathbf{x}^M - \mathbf{y}^M||^2 = \sum_{m=1}^{M} ((A\mathbf{x}^{m-1})_m - (A\mathbf{y}^{m-1})_m)^2.$$

Solution: Calculate $||\mathbf{x}^m - \mathbf{y}^m||^2 - ||\mathbf{x}^{m+1} - \mathbf{y}^{m+1}||^2$ for each $m = 0, 1, ..., M-1$ and then add.

Exercise 50.3: Show that if we select B so that C is invertible and $B^T A = 0$ then the exact solution of $C\mathbf{z} = \mathbf{b}$ is the concatenation of the least squares solutions of $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{y} = \mathbf{b}$.

Solution: Calculate the solution of $C\mathbf{z} = \mathbf{b}$ as the least squares solution of $C\mathbf{z} = \mathbf{b}$.

Chapter 56: The Wave Equation

Exercise 1: Show that the radial function $u(r,t) = \frac{1}{r}h(r-ct)$ satisfies the wave equation for any twice differentiable function h.

Solution: The partial derivatives are as follows:

$$u_t = -c\frac{1}{r}h'(r-ct),$$

$$u_{tt} = c^2\frac{1}{r}h''(r-ct),$$

$$u_r = -\frac{1}{r^2}h(r-ct) + \frac{1}{r}h'(r-ct),$$

and

$$u_{rr} = 2\frac{1}{r^3}h(r-ct) - \frac{2}{r^2}h'(r-ct) + \frac{1}{r}h''(r-ct).$$

The result follows immediately from these facts.