6.1.1. Approximation of Functions

Let $v(t)$ be a given function to be approximated using well-known functions $w_1(t), \ldots, w_n(t)$. One approach is to seek coefficients $c_1, \ldots, c_n$ that minimize

$$\left\| v - \sum_{j=1}^{n} c_j w_j \right\|^2 = \int \left| v(t) - \sum_{j=1}^{n} c_j w_j(t) \right|^2 dt.$$

Let $W := \operatorname{span}\{w_1, \ldots, w_n\}$. Then optimal coefficients are given by (6.11), where

$$G_{ij} = \int w_j(t) w_i(t)\, dt \quad \text{and} \quad b_i = \int v(t) \overline{w_i(t)}\, dt.$$

Problem 6–16. You wish to approximate the function $v(t) = t^3$ for $0 \le t \le 1$ using a polynomial of degree 1; i.e., $\widehat{v}(t) = c_1 + c_2 t$. Find numerical values of c_1 and c_2 that minimize the mean squared error,

$$\int_0^1 \left| v(t) - \widehat{v}(t) \right|^2 dt.$$

Problem 6–17. Find the best approximation (in the sense of mean squared error) of $v(t) = t^2, 0 \le t \le 1$, by a polynomial $\widehat{v}(t)$ of degree at most one that also satisfies $\int_0^1 \widehat{v}(t)\, dt = 0$.

If $v(t)$ is not given theoretically, or if the integrals for b_i above cannot be computed numerically, then the foregoing approach is not possible. However, suppose samples $v(t_1), \ldots, v(t_m)$ are available. Put

$$\mathbf{v} := [v(t_1), \ldots, v(t_m)]' \quad \text{and} \quad \mathbf{w}_j := [w_j(t_1), \ldots, w_j(t_m)]'.$$

Using the standard inner product on $\mathbb{R}^n$ or $\mathbb{C}^n$ and the corresponding Euclidean distance, we have

$$\left\| \mathbf{v} - \sum_{j=1}^{n} c_j \mathbf{w}_j \right\|^2 = \sum_{i=1}^{m} \left| v(t_i) - \sum_{j=1}^{n} c_j w_j(t_i) \right|^2. \tag{6.18}$$

If we put $\mathscr{W} := [\mathbf{w}_1, \ldots, \mathbf{w}_n]$, then in $Gc = b$, $G = \mathscr{W}^H \mathscr{W}$, and $b = \mathscr{W}^H \mathbf{v}$. Note that once we have found the coefficients $c_1, \ldots, c_n$, the approximation

$$\widehat{v}(t) = \sum_{j=1}^{n} c_j w_j(t)$$

can be evaluated for all t for which the $w_j(t)$ are defined, not just the sample times $t_1, \ldots, t_m$. Hence, for $t_i < t < t_{i+1}$, $\widehat{v}(t)$ can serve as an approximation of $v(t)$.

Remark. Suppose that the $\mathbf{w}_j$ are linearly independent and that $m = n$. Then the $\mathbf{w}_j$ are a basis for n-dimensional Euclidean space, and the minimum value of (6.18) is zero when the c_j are the unique coefficients of the representation of $\mathbf{v}$ in this basis. In other words, the approximation $\widehat{v}(t)$ satisfies $\widehat{v}(t_i) = v(t_i)$ for all i; i.e., the approximation **interpolates** the data.

Application (Polynomial Approximation). We can compute polynomial approximations in the sense of (6.18) very easily using MATLAB. For example, suppose we want to approximate $v(t) = \sin(2\pi t)$ for $t \in [0,1]$ using a polynomial of degree 4. This corresponds to projecting v onto the 5-dimensional space spanned by $1, t, t^2, t^3, t^4$. Use the following commands.

```
t = linspace(0,1,30);
v = sin(2*pi*t);
c = polyfit(t,v,4); % 4th deg polynomial
tt = linspace(0,1,100);
vhat = polyval(c,tt);
plot(t,v,'-o',tt,vhat,'--')
```

In polynomial approximation, if $m = n+1$, then the approximation will interpolate the data.

Theorem 7.11. *Let X and Y be vector spaces, and let $A\colon X \to Y$ be a linear operator. If* $\dim X < \infty$*, then both* $\ker A$ *and* $\operatorname{range} A$ *are finite dimensional, and*

$$\dim \ker A + \dim \operatorname{range} A \;=\; \dim X.$$

Proof. By Problem 1–3, $\dim \ker A \le \dim X$. Put $n := \dim X$, and put $r := \dim \ker A$. Let $\{x_1, \ldots, x_r\}$ be a basis for $\ker A$. Extend this to a basis for X, say $\{x_1, \ldots, x_r, x_{r+1}, \ldots, x_n\}$. Then it is easy to show that $\{Ax_{r+1}, \ldots, Ax_n\}$ is a basis for $\operatorname{range} A$; i.e., $\dim \operatorname{range} A = n - r = \dim X - \dim \ker A$. □

Problem 7–14. Show that $\{Ax_{r+1}, \ldots, Ax_n\}$ in the preceding proof is a basis for $\operatorname{range} A$.

Problem 7–15. Let X and Y be vector spaces, and let $A\colon X \to Y$ be a linear operator. Suppose $\dim X < \infty$ and A is invertible. Show that $\dim X = \dim Y$.

Recall from Problems 2–2 and 2–3 that to show that a function is invertible, it is usually necessary to show that it is both one-to-one *and* onto. The next proposition covers an important special case for which it is enough to prove just one of these properties, as the other holds automatically!

Proposition 7.12. *Let X and Y be finite-dimensional vector spaces, and let $A\colon X \to Y$ be a linear operator. If* $\dim X = \dim Y$*, then A is nonsingular if and only if A is onto.*

Proof. Recall that $\operatorname{range} A$ is a subspace of Y. If we can show that $\dim \operatorname{range} A = \dim Y$, then by Problem 1–3, $\operatorname{range} A = Y$. Suppose that A is nonsingular. Then $\dim \ker A = 0$, and using Theorem 7.11, we can obtain $\dim \operatorname{range} A = \dim X = \dim Y$. Conversely, if A is onto, then $\operatorname{range} A = Y$. This implies that $\dim \operatorname{range} A = \dim Y = \dim X$. Combining this with Theorem 7.11 yields $\dim \ker A = 0$; hence, $\ker A$ is the zero subspace, and A is nonsingular. □

The situation in Theorem 7.11, in which A maps a finite-dimensional space into an infinite-dimensional space, arises frequently in communication systems.

Example 7.13 (Modulation Operator). Let $\varphi_1(t), \ldots, \varphi_n(t)$ be finite-energy **signaling waveforms**. For $x = [x_1, \ldots, x_n]' \in \mathbb{C}^n$, put

$$(Ax)(t) \;=\; \sum_{k=1}^{n} x_k \varphi_k(t). \tag{7.2}$$

Thus, $A\colon \mathbb{C}^n \to L^2[0,T]$. This operator can be implemented as shown in in Figure 5.

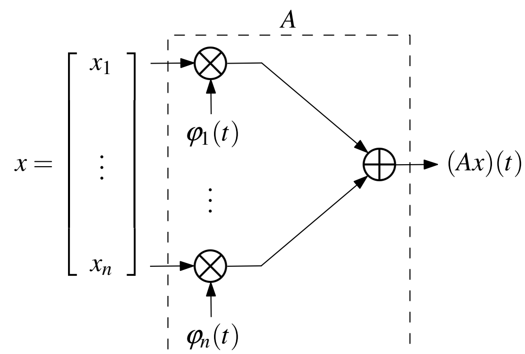

Figure 5. Block diagram for the "modulation operator" A in (7.2).

Example 7.19. Let A denote the "modulation operator" of Example 7.13. Let $\langle\cdot,\cdot\rangle$ denote the inner product on $L^2[0,T]$. We can find the adjoint by inspection as follows. First note that since $A\colon \mathbb{C}^n \to L^2[0,T]$, $A^*\colon L^2[0,T] \to \mathbb{C}^n$. Hence, the formula we are looking for must be such that A^*y is an n-dimensional column vector. Write

$$\begin{aligned}
\langle Ax, y\rangle &= \int_0^T (Ax)(t)\overline{y(t)}\,dt \\
&= \int_0^T \left[\sum_{k=1}^n x_k\varphi_k(t)\right]\overline{y(t)}\,dt \\
&= \sum_{k=1}^n x_k \overline{\int_0^T y(t)\overline{\varphi_k(t)}\,dt} \\
&= \sum_{k=1}^n x_k \overline{\langle y, \varphi_k\rangle}.
\end{aligned}$$

This last expression is the Euclidean inner product of $x = [x_1,\ldots,x_n]'$ with the column vector whose kth component $\langle y, \varphi_k\rangle$. Hence,

$$A^*y = \begin{bmatrix} \langle y,\varphi_1\rangle \\ \vdots \\ \langle y,\varphi_n\rangle \end{bmatrix}.$$

This adjoint operator can be implemented as shown in Figure 6.

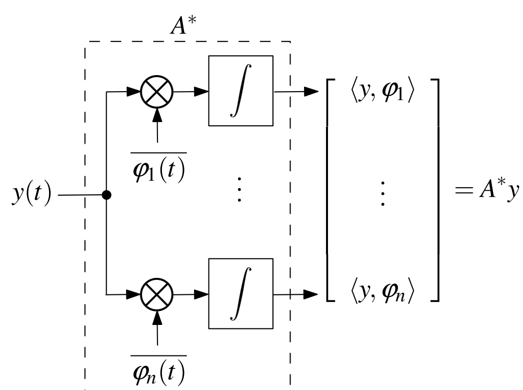

Figure 6. Implementation of adjoint operator.

Remark. An inner product of the form $\int_0^T y(t)\overline{\varphi(t)}\,dt$ can always be expressed as a sampled convolution with impulse response $h(\theta) = \overline{\varphi(T-\theta)}$. To see this write

$$\begin{aligned}
\left(\int_{-\infty}^{\infty} h(t-\tau)y(\tau)\,d\tau\right)\bigg|_{t=T} &= \left(\int_{-\infty}^{\infty} h(T-\tau)y(\tau)\,d\tau\right) \\
&= \left(\int_{-\infty}^{\infty} \overline{\varphi(T-[T-\tau])}y(\tau)\,d\tau\right) \\
&= \langle y, \varphi\rangle,
\end{aligned}$$

where we have assumed that $\varphi(\tau) = 0$ for τ outside $[0,T]$. Because the impulse response h is defined in terms of the signal φ, h is said to be "matched" to the signal. Hence, h is called a **matched filter**. Letting $H_k(f)$ denote the Fourier transform of $\overline{\varphi_k(T-t)}$, we see that Figure 6 can also be viewed as the bank of matched filters in Figure 7.

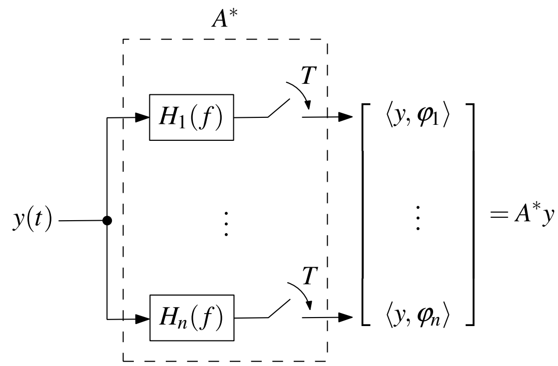

Figure 7. Matched filter equivalent of Figure 6.

Problem 7–34. Assume A^* exists.

(a) Show that $\ker A^* = (\operatorname{range} A)^\perp$.
(b) Show that $(A^*)^* = A$.
(c) Show that $\ker A = (\operatorname{range} A^*)^\perp$.
(d) Show that $(\ker A)^\perp \supset \operatorname{range} A^*$.
Hint: Problem 6–11(c).
(e) If X is a Hilbert space, and if $A{:}X \to Y$ is a bounded linear operator, show that $(\ker A)^\perp = \overline{\operatorname{range} A^*}$.
Hint: Use Problem 6–24.
(f) Show that $\ker A^*A = \ker A$.

Example 7.24 (Digital Communication Systems). Consider a communication system employing the "modulation operator" A defined in Example 7.13 and illustrated in Figure 5. We claim that a reasonable receiver structure begins with the operator A^* illustrated in Figure 6. If we transmit the waveform $y = Ax$, and the receiver produces $A^*y = A^*Ax$, then x can be recovered from A^*y by computing

$$(A^*A)^{-1}(A^*y) \;=\; (A^*A)^{-1}(A^*A)x \;=\; x.$$

Such a system is shown in Figure 8. The point is that since

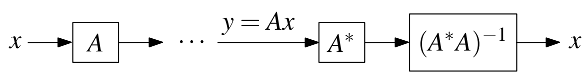

Figure 8. An ideal, noiseless communication system.

the φ_k in (7.2) are linearly independent, A is nonsingular, and by Problem 7–34(f), so is A^*A. Furthermore, since A^*A maps $\mathbb{C}^n$ into itself, we have from Proposition 7.12 that A^*A is invertible. Of course, receiver processing is greatly simplified if A^*A is diagonal. This is the case in orthogonal frequency division multiplexing (OFDM), in which

$$\varphi_k(t) \;=\; e^{j2\pi(k/T)t}, \quad 0 \le t \le T.$$

Remark. The foregoing example can be slightly generalized. Before transmitting a vector $x \in \mathbb{C}^n$, first apply an invertible $n \times n$ matrix W; i.e., transmit $y = A(Wx)$. The receiver computes A^*y as before. Now observe that

$$W^{-1}(A^*A)^{-1}(A^*y) \;=\; W^{-1}(A^*A)^{-1}(A^*A)Wx \;=\; x.$$

Problem 7–35. Let A and A^* be as in Example 7.24. For $x = [x_1, \ldots, x_n]$, show that the ith component of $A^*(Ax)$ is given by

$$\sum_{k=1}^{n} \langle \varphi_k, \varphi_i \rangle x_k.$$

Since $(A^*A)x$ can be computed by applying the matrix with ik entry $\langle \varphi_k, \varphi_i \rangle$ to the column vector x, the operator $(A^*A)^{-1}$ can be implemented by applying the inverse of this matrix.

If A^*A is invertible, then (7.5) is equivalent to

$$x_0 = (A^*A)^{-1}A^*y_0.$$

By Problem 7–34(f), A^*A is nonsingular if and only if A is nonsingular. If X is finite dimensional, then A^*A being nonsingular implies that it is onto by Proposition 7.12, and therefore invertible.

Remark. When A^*A is invertible, we have a formula for the projection of y_0 onto range$A$, namely (cf. (6.17)),

$$\widehat{y_0} = Ax_0 = A(A^*A)^{-1}A^*y_0. \tag{7.6}$$

This expresses the projection as a function of A^*y_0. Applying A^* to (7.6) shows that A^*y_0 is a function of the projection. Hence, A^*y_0 and the projection of y_0 onto range$A$ *contain the same information*.

Remark (Continued). In the digital communication scenario of Example 7.24, it is easier to work with A^*y than the projection of y because A^*y is a column vector and the projection of y is a waveform. But what really motivates the receiver design is the idea of projecting the received waveform onto the subspace spanned by the signaling waveforms. Even though Ax is transmitted, the signal at the receiver is $Ax+z$, where z is a noise waveform. Since Ax is in the range of A, the projection of $Ax+z$ is equal to the sum of Ax and the projection of z onto the range of A; i.e., there is no loss of information about the transmitted signal Ax. There is the additional benefit that the energy of the projected noise waveform is no greater than that of the noise waveform itself (recall (6.8)).[10]

[10] Although the receiver does not lose any information about the signal by doing the projection, the reader may wonder if the receiver loses information about the noise that could be helpful. If the noise is white and Gaussian, it can be proved that nothing is lost. Otherwise, projection can be suboptimal. Consider the received vector

$$x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} z \\ -z \end{bmatrix}.$$

If we project onto the space spanned by $[1,0]'$, we get $[x+z,0]'$ and lose the information in the second dimension. However, if we add the first and second components we recover x without noise.