Chapter 11

Kalman and Wiener Filtering

In this chapter we will present a brief introduction to Kalman and Wiener filtering. The main emphasis is to develop a connection between filtering theory and the Riccati equation.

Let us briefly review the notion of random variable. Consider a probability space $(\Omega, \mathcal{A}, P)$ where $\Omega$ is the universal set, $\mathcal{A}$ is a $\sigma$-algebra, and $P$ is the probability measure. Recall that a random variable $x$ is a measurable function mapping $\mathcal{A}$ into $\mathbb{C}$. The mean of $x$ is given by

$$E_x = \int_{\Omega} x dP$$

where $E$ denotes the expectation. The variance of $x$ is determined by

$$\sigma_x^2 = E|x - \mu_x|^2 = \int_{\Omega} |x - \mu_x|^2 dP.$$  

Moreover, $\sigma_x$ is the standard deviation for $x$. Let $f$ and $g$ be two measurable functions of random variables $x$ and $y$ respectively. If $x$ and $y$ are independent, then it follows from probability theory that $Ef(x)g(y) = Ef(x)Eg(y)$.

We say that a sequence $\{y(n)\}_{n=\infty}^{-\infty}$ is a stochastic process or random process if each $y(n)$ is a random variable. Let $L^2(\Omega, \mathcal{A}, P)$ be the Hilbert space of all square integrable random variables with respect to probability measure $dP$. The inner product is given by $(x, y) = Exy$ for all $x$ and $y$ in $L^2(\Omega, \mathcal{A}, P)$. Throughout we assume that all random variables are in $L^2(\Omega, \mathcal{A}, P)$. Moreover, let $\{u(n)\}$ and $\{v(n)\}$ be random processes where each $u(n)$ and $v(n)$ is an element in $L^2(\Omega, \mathcal{A}, P)$. We say that the random processes $u(n)$ and $v(n)$ are orthogonal, if for all integers $i$ and $j$, the random variable $u(i)$ is orthogonal to $v(j)$, or equivalently, the inner product $(u(i), v(j)) = Eu(i)v(j) = 0$. In this case, let $y(n)$ be a random process given by $y(n) = u(n) + v(n)$. Then we have

$$(y(i), y(j)) = (u(i), u(j)) + (v(i), v(j)).$$
Finally, if \( u(n) \) and \( v(n) \) are independent random processes and \( u(n) \) or \( v(n) \) has zero mean for all \( n \), then \( u(n) \) and \( v(n) \) are orthogonal. To verify this, simply observe that
\[
(u(i), v(j)) = Eu(i)v(j) = Eu(i)Ev(j) = 0
\]
for all \( i \) and \( j \). Therefore \( u(n) \) and \( v(n) \) are orthogonal.

### 11.1 Random Vectors

Recall that \( E \) denotes the expectation. In particular, \( Eg \) is the mean of the random variable \( g \). Let \( K \) be the Hilbert space generated by the set of all random variables \( g \) such that \( E|g|^2 \) is finite. Throughout we always assume that all of our random variables are in \( K \). The inner product on \( K \) is determined by the expectation, that is, \((f, g) = Ef\overline{g}\) where \( f \) and \( g \) are in \( K \). We say that \( f \) is a random vector with values in \( \mathbb{C}^k \) if \( f \) is a vector of the form \( f = [f_1 \ f_2 \ \cdots \ f_k]^\text{tr} \) where \( \{f_j\}_1^k \) are all random variables. (Recall that \( \text{tr} \) denotes the transpose.) In this case, \( Ef \) is the vector in \( \mathbb{C}^k \) defined by
\[
Ef = [Ef_1 \ Ef_2 \ \cdots \ Ef_k]^\text{tr}.
\]
The correlation matrix \( R_f \) is the matrix on \( \mathbb{C}^k \) defined by \( R_f = Ef f^* \). To be precise,
\[
R_f = Ef f^* = \begin{pmatrix}
Ef_1 \overline{f}_1 & Ef_1 \overline{f}_2 & \cdots & Ef_1 \overline{f}_k \\
Ef_2 \overline{f}_1 & Ef_2 \overline{f}_2 & \cdots & Ef_2 \overline{f}_k \\
\vdots & \vdots & \ddots & \vdots \\
Ef_k \overline{f}_1 & Ef_k \overline{f}_2 & \cdots & Ef_k \overline{f}_k
\end{pmatrix}.
\] (11.1.1)

Notice that the \( j-k \) entry of \( R_f \) is given by \((R_f)_{jk} = Ef_j \overline{f}_k\). Finally, it is noted that \( R_f \) is the Gram matrix determined by \( \{f_j\}_1^k \). The following result shows that \( R_f \) is positive.

**Theorem 11.1.1.** Let \( f = [f_1 \ f_2 \ \cdots \ f_k]^\text{tr} \) be a random vector with values in \( \mathbb{C}^k \). Then \( R_f \) is a positive matrix on \( \mathbb{C}^k \). Moreover, \( R_f \) is strictly positive (\( R_f > 0 \)) if and only if the random variables \( \{f_j\}_1^k \) are linearly independent.

**Proof.** Let \( \alpha = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_k]^\text{tr} \) be any vector in \( \mathbb{C}^k \). Then
\[
(R_f \alpha, \alpha) = (Ef f^* \alpha, \alpha) = Ef f^* \alpha = E\|f^* \alpha\|^2 = E \left| \sum_{j=1}^k \overline{f}_j \alpha_j \right|^2 \geq 0. \tag{11.1.2}
\]
Hence \((R_f \alpha, \alpha) \geq 0\) for all \( \alpha \) in \( \mathbb{C}^k \). Therefore \( R_f \) is positive.

Equation (11.1.2) shows that
\[
(R_f \alpha, \alpha) = E \left| \sum_{j=1}^k \overline{f}_j \alpha_j \right|^2.
\]