Solution Set 4

33. *i)* Let us first verify the result given in the *hint*: For any $f, g \in H$,

$$\begin{aligned}\|f+g\|^2 - \|f-g\|^2 + \imath\|f+\imath g\|^2 - \imath\|f-\imath g\|^2 &= 2(f,g) + 2(g,f) + 2\imath(f,\imath g) + 2\imath(\imath g, f) \\ &= 4(f,g),\end{aligned}$$

where the first equality has followed by simply writing the norms in terms of inner products, and then expanding the inner products and cancelling out some terms. The second equality follows from the facts that

$$(f,\imath g) = -\imath(f,g) \quad \text{and} \quad (\imath g, f) = \imath(g,f)$$

Now since $A = B$, $\sum |(f,\xi_j)|^2 = A\,\|f\|^2$, and hence

$$\begin{aligned}&\sum_{j\in J} \{|(f+g,\xi_j)|^2 - |(f-g,\xi_j)|^2 + \imath|(f+\imath g,\xi_j)|^2 - \imath|(f-\imath g,\xi_j)|^2\} \\ &\qquad = A\{\|f+g\|^2 - \|f-g\|^2 + \imath\|f+\imath g\|^2 - \imath\|f-\imath g\|^2\}\end{aligned}$$

From the *hint* just verified, the RHS is equal to $4(f,g)$; the LHS, on the other hand, simplifies to (by simply multiplying out the inner products): $4\sum_{j\in J}(f,\xi_j)\,(\xi_j,g)$. This leads to the desired result, since LHS = RHS for all $g \in H$.

ii) Let $f = (v_1, v_2)'$, where $v_i \in \mathbf{C}$, $i = 1, 2$. Then,

$$\begin{aligned}|(f,\xi_1)|^2 + |(f,\xi_2)|^2 + |(f,\xi_3)|^2 &= |v_2|^2 + |\frac{\sqrt{3}}{2}v_1 + \frac{1}{2}v_2|^2 + |\frac{\sqrt{3}}{2}v_1 - \frac{1}{2}v_2|^2 \\ &= \frac{3}{2}(|v_1|^2 + |v_2|^2) = \frac{3}{2}\|f\|^2\end{aligned}$$

The frame bound is 3/2. Note that the given set is linearly dependent.

iii) Clearly,

$$(f,\xi_j) = 0 \;\forall\, j \quad \Rightarrow \quad \|f\| = 0 \quad \Rightarrow \quad f \equiv 0$$

Therefore, $\{\xi_j\}$ span all of H. To show that it is an orthonormal set, in the equality of *part i)* pick $f = \xi_k$ for an arbitrary $k \in J$, and take inner product of both sides with ξ_k:

$$\|\xi_k\|^2 = \sum_{j\in J} |(\xi_k,\xi_j)|^2 = \|\xi_k\|^4 + \sum_{j\neq k} |(\xi_k,\xi_j)|^2$$

Since, by hypothesis, $\|\xi_k\| = 1$, the equality above can hold iff $(\xi_k,\xi_j) = 0 \;\forall\, j \neq k$. Since k was arbitrary, the desired result follows.

34. *i)* Let

$$M = \left\{ z \in L_2(\Omega, \mathcal{P},\ \mathbf{R}^n) : z = \sum_{j=0}^{i} K_j y_j;\ \ K_j \in \mathcal{M}_{nm} \right\}.$$

This is a closed linear subspace of $L_2(\Omega, \mathcal{P};\ \mathbf{R}^n)$ which is a Hilbert space. Hence, from the Projection Theorem, there exists a unique $\hat{x} \in M$, expressed as $\hat{x} = \sum_{j=0}^{i} \hat{K}_j\, y_j$, with

$$\inf_{\substack{K_j \in \mathcal{M}_{nm} \\ j=0,\ldots,i}} \|x - \sum_{j=0}^{i} K_j y_j\| = \|x - \hat{x}\|$$

and a necessary and sufficient condition for $\hat{x}$ to be the minimizing solution is

$$(x - \hat{x}, z) = 0 \qquad \forall z \in M$$

$\Leftrightarrow$

$$E[x^T Q K_\ell y_\ell] = \sum_{j=0}^{i} E\left[y_j^T \hat{K}_j^T Q K_\ell y_\ell\right] \quad \forall K_\ell \in \mathcal{M}_{nm} \quad \ell = 0, 1, \ldots, i$$

$\Leftrightarrow$

$$\text{Tr}\ \left\{E\left[y_\ell x^T Q K_\ell\right]\right\} = \sum_{j=0}^{i}\ \text{Tr}\ \left\{E\left[y_\ell y_j^T \hat{K}_j^T Q K_\ell\right]\right\} \quad \forall K_\ell \in \mathcal{M}_{nm}.$$

Since Q is positive definite, $QK_\ell \in \mathcal{M}_{nm}$ whenever $K_\ell \in \mathcal{M}_{nm}$, and vice versa, and hence the earlier condition is

$$\text{Tr}\ \left\{\left(\Lambda_{\ell x} - \sum_{j=0}^{i} \Lambda_{\ell j} \hat{K}_j^T\right) K\right\} = 0 \quad \forall K \in \mathcal{M}_{nm} \quad \ell = 0, 1, \ldots, i$$

where

$$\boxed{\Lambda_{\ell x} \triangleq E\left[y_\ell x^T\right] \ ; \quad \Lambda_{\ell j} \triangleq E[y_\ell y_j^T]}.$$

Now, two random vectors are uncorrelated if their components (considered as random variables) are uncorrelated. Furthermore, since $E[y_\ell] = 0$, we have

$$\Lambda_{\ell j} = \begin{cases} \Lambda_{\ell\ell} & j = \ell \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the condition now becomes

$$\text{Tr}\ \left\{\left(\Lambda_{\ell x} - \Lambda_{\ell\ell} \hat{K}_\ell^T\right) K\right\} = 0 \qquad \forall K \in \mathcal{M}_{nm},\ \ \ell = 0, \ldots, i;$$

and assuming that $\Lambda_{\ell\ell}$ is invertible ($\ell = 0, \ldots, i$), we have a unique solution

$$\boxed{\hat{K}_\ell = \Lambda_{\ell x}^T \Lambda_{\ell\ell}^{-1}, \quad \ell = 0, 1, \ldots, i.}$$

Note: Uniqueness of the $\hat{K}_j$'s follows from the fact that $\text{Tr}\ [AB] = 0\ \forall B \Rightarrow A$ is a matrix with only zero entries (i.e., the zero matrix).

If $\Lambda_{\ell\ell}$ is singular, then the solution is still unique in M (which follows from the *Projection Theorem*, as stated earlier), but the corresponding $\hat{K}_j$'s may not be unique. They will then be solved from the equations:

$$\hat{K}_\ell \Lambda_{\ell\ell} = \Lambda_{\ell x}^T,\ \ell = 0, 1, \ldots, i.$$

ii)

$$
\begin{aligned}
\epsilon_k &= \left\| x - \sum_{j=0}^{k-1} \hat{K}_j y_j - \hat{K}_k y_k \right\|^2 = \underbrace{\left\| x - \sum_{j=0}^{k-1} \hat{K}_j y_j \right\|^2}_{\epsilon_{k-1}} + \|\hat{K}_k y_k\|^2 - 2\left(x - \sum_{j=0}^{k-1} \hat{K}_j y_j , \hat{K}_k y_k \right) \\
&= \epsilon_{k-1} + \text{ Tr } \left[\Lambda_{kk} \hat{K}_k^T Q \hat{K}_k \right] - 2 \text{ Tr } \left\{ E\left[y_k x^T Q \hat{K}_k \right] \right\} \\
&= \epsilon_{k-1} + \text{ Tr } \left[\Lambda_{kx} Q \Lambda_{kx}^T \Lambda_{kk}^{-1} \right] - 2 \text{ Tr } \left\{ \Lambda_{kx} Q \Lambda_{kx}^T \Lambda_{kk}^{-1} \right\}
\end{aligned}
$$

$\Leftrightarrow$

$$
\boxed{\epsilon_k = \epsilon_{k-1} - \text{ Tr } \left[\Lambda_{kx} Q \Lambda_{kx}^T \Lambda_{kk}^{-1} \right].}
$$

35. *i)* Note that $\mathcal{Z}$ is a closed, convex subset of the Hilbert space $L_2(\Omega, P;\ \mathbf{R})$ (*closed*, because $a_1 \geq 0, a_2 \geq 0$, and *convex*, because the set of $(a_1, a_2) \in \mathbf{R}^2$, $a_1 \geq 0, a_2 \geq 0$, is convex). Then the problem is one of minimizing $\|x - z\|$ over $z \in \mathcal{Z}$, or (equivalently) one of finding the minimum distance from $x \in L_2(\Omega, P;\ \mathbf{R})$ to $\mathcal{Z}$.

ii) It follows from Theorem 1, p. 69 of *Luenberger* that there exists a unique $\hat{x} \in \mathcal{Z}$ that minimizes $\|x - z\|$ on $\mathcal{Z}$. Furthermore,

$$
(x - \hat{x}, z - \hat{x}) \leq 0 \quad \forall\, z \in \mathcal{Z}
$$

iii) Writing $\hat{x}$ and $z \in \mathcal{Z}$ as $\hat{x} = \hat{a}_1 y_1 + \hat{a}_2 y_2$ and $z = a_1 y_1 + a_2 y_2$, the condition in part (ii) above becomes

$$
\begin{aligned}
(x - \hat{a}_1 y_1 - \hat{a}_2 y_2, (a_1 - \hat{a}_1) y_1 + (a_2 - \hat{a}_2) y_2) &= (0.2 - \hat{a}_1)(a_1 - \hat{a}_1) - (0.5 + \hat{a}_2)(a_2 - \hat{a}_2) \\
&\leq 0 \quad \forall\, a_1 \geq 0, a_2 \geq 0
\end{aligned}
$$

Pick $a_1 = \hat{a}_1 \quad \Rightarrow \quad (0.5 + \hat{a}_2)(a_2 - \hat{a}_2) \geq 0 \ \forall\, a_2 \geq 0 \quad \Leftrightarrow \quad \hat{a}_2 = 0$

Now pick $a_2 = \hat{a}_2 = 0 \quad \Rightarrow \quad (0.2 - \hat{a}_1)(a_1 - \hat{a}_1) \leq 0 \ \forall\, a_1 \geq 0 \quad \Leftrightarrow \quad \hat{a}_1 = 0.2$. Hence,

$$
\hat{x} = 0.2\, y_1\,, \quad E[(x - \hat{x})^2] = \|x - \hat{x}\|^2 = \|x - 0.2 y_1\|^2 = 0.96
$$

36. Let H be the Hilbert space of second-order random variables defined on $(\Omega, \mathcal{F}, \mathcal{P})$. Then, this problem can be posed as a *minimum norm* problem on H, subject to equality constraints. That is, letting $(\cdot, \cdot)$ denote the norm on H,

$$
\min_{x \in H} \|x\| \quad \text{such that} \quad (x, 1) = -1 \ \text{ and } \ (x, y) = 2
$$

i) Note that y and 1 are linearly independent on H, because if they were linearly dependent, then we would have from $E[y] = 1$, $y = 1$ with probability 1, which however contradicts with $E[y^2] = 2$. Then, from the *Projection Theorem* the problem admits a unique solution in $M = [1, y]$.

ii) Since the solution is in M, it is in the form $x^\circ = \alpha_1 + \alpha_2 y$, for two real numbers α_1 and α_2. These can be obtained from the given two constraints:

$$(\alpha_1+\alpha_2 y, 1) = -1, (\alpha_1+\alpha_2 y, y) = 2 \quad \Rightarrow \quad \alpha_1+\alpha_2 = -1, \alpha_1+2\alpha_2 = 2 \quad \Rightarrow \quad \alpha_1 = -4, \alpha_2 = 3$$

iii) Replace the constraint $(x, 1) = -1$ above with $(x, 1) = c$ where $c \geq -1$. For each fixed c, the problem again admits a unique solution (by the same reasoning as above), and is given by

$$x^\circ = \alpha_1 + \alpha_2 y\,, \quad \alpha_1 + \alpha_2 = c, \alpha_1 + 2\alpha_2 = 2 \quad \Rightarrow \quad \alpha_1 = 2(c-1), \alpha_2 = 2-c$$

Then, $\|x^\circ\|^2 = 4(c-1)^2 4(c-1)(2-c) + 2(2-c)^2$, which is minimized uniquely by choosing $c = 1$ (simply find the stationary point of the strictly convex function $\|x^\circ\|^2$). This then leads to

$$\alpha_1 = 0\,, \; \alpha_2 = 1 \quad \Rightarrow \quad x^\circ = y$$

37. Let $E[Y_i] \triangleq \bar{y}_i$, $E[Y_i^2] = r_i$, and take w to be a positive function in $C[0, 1]$. Let $\tilde{k}_i(t) = k_i(t)/w(t)$. k_1 and k_2 linearly independent in $C[0, 1] \Rightarrow \tilde{k}_1$ and $\tilde{k}_2$ are linearly independent in $L_2[0, 1]$. Then, the optimization problem is:

$$\min \|X\|; \quad X \in L_2(\Omega, \mathcal{P}; C[0, 1])$$

such that

$$(X, \tilde{k}_1 Y_1) = c_1; \quad (X, \tilde{k}_2 Y_2) = c_2.$$

Let us first consider the case where X is allowed to be in $L_2(\Omega, \mathcal{P}; L_2[0, 1])$. Since this is a Hilbert space, the conditions of Theorem 2, p. 65 of the text, are satisfied, implying that the solution to the second problem, say $\hat{x}$, is unique and is given by

$$\hat{x}(t, w) = \beta_1 \tilde{k}_1(t) Y_1(\omega) + \beta_2 \tilde{k}_2(t) Y_2(\omega)$$

where

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \|\tilde{k}_1 Y_1\|^2 & (\tilde{k}_1 Y_1, \tilde{k}_2 Y_2 \\ (\tilde{k}_2 Y_2, \tilde{k}_1 Y_1) & \|\tilde{k}_2 Y_2\|^2 \end{pmatrix}^{-1} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

The matrix here is invertible, because $\tilde{k}_1$ and $\tilde{k}_2$ are linearly independent (WHY ?). Since Y_1 and Y_2 are uncorrelated

$$\begin{aligned} (\tilde{k}_1 Y_1, \tilde{k}_2 Y_2) &= (\tilde{k}_2 Y_2, \tilde{k}_1 Y_1) = E[\int_0^1 \frac{k_1(t)}{w(t)} k_2(t)\, dt\, Y_1 Y_2] \\ &= K_{12} \bar{y}_1 \bar{y}_2, \text{ where } K_{12} = \int_0^1 \frac{k_1(t) k_2(t)}{w(t)} dt. \end{aligned}$$

Furthermore,

$$\|\tilde{k}_i Y_i\|^2 \quad = \quad r_i \int_0^1 [k_i^2(t)/w(t)]\, dt \triangleq r_i K_i$$

$$\begin{pmatrix} r_1 K_1 & K_{12}\bar{y}_1\bar{y}_2 \\ K_{12}\bar{y}_1\bar{y}_2 & r_2 K_2 \end{pmatrix}^{-1} = \frac{1}{r_1 r_2 K_1 K_2 - (K_{12}\bar{y}_1\bar{y}_2)^2} \begin{pmatrix} r_2 K_1 & -K_{12}\bar{y}_1\bar{y}_2 \\ -K_{12}\bar{y}_1\bar{y}_2 & r_1 K_1 \end{pmatrix}$$

$\therefore$

$$\begin{aligned} \hat{\beta}_1 &= (c_1 r_2 K_2 - c_2 K_{12}\bar{y}_1\bar{y}_2)/[r_1 r_2 K_1 K_2 - (K_{12}\bar{y}_1\bar{y}_2)^2] \\ \hat{\beta}_2 &= (c_2 r_1 K_1 - c_1 K_{12}\bar{y}_1\bar{y}_2)/[r_1 r_2 K_1 K_2 - (K_{12}\bar{y}_1\bar{y}_2)^2] \\ \hat{x}(t,w) &= \hat{\beta}_1 \tilde{k}_1(t) Y_1(\omega) + \hat{\beta}_2 \tilde{k}_2(t) Y_2(\omega) \end{aligned}$$

Note that $\hat{x}$ also belongs to $L_2(\Omega, \mathcal{P}; C[0,1])$, because $\tilde{k}_1$ and $\tilde{k}_2$ are continuous. Hence, $\hat{x}$ is the unique vector solving the original problem.

38. i) First note that m as defined is a second-order random variable if $K(\cdot)$ is square-integrable, because by the Cauchy-Schwartz inequality,

$$E[m^2] \leq \int_0^2 |K(t)|^2 dt \int_0^2 E[Y^2(t)] dt$$

and the second product term is finite. Now, the set of all m's corresponding to such K's is a closed subspace M of $L_2(\Omega, \mathcal{P};\ \mathbf{R})$, and hence by the Projection Theorem, there is a unique projection of X onto M, and by orthogonality we have the equation

$$R_{XY}(t) = \int_0^2 \hat{K}(s) R_{YY}(s,t) ds$$

which has a solution $\hat{K}$, and in terms of this,

$$\hat{X} = \int_0^2 \hat{K}(t) Y(t) dt$$

is the unique element of M. And $\hat{K}$ is unique if ${}_{YY}(s,t)$ is positive definite.

ii) Now M is the one-dimensional subspace generated by the random variable

$$Z(\omega) := \int_0^2 Y(t;\omega) dt\,.$$

Note that this random variable has second moment (or equivalently norm squared)

$$\sigma_Z^2 := \int_0^2 \int_0^2 R_{YY}(s,t)\, ds\, dt\,,$$

and

$$E[XZ] = \int_0^2 R_{XY}(t)dt =: \sigma_{XZ} \,.$$

Now we have projection of one random variable (X) into the subspace generated by another (Z), and the result follows from standard theory discussed in class:

$$\hat{X} = (\sigma_{XZ}/\sigma_Z^2)Z$$

where we have assumed without any loss of generality that σ_Z^2 is nonzero, because otherwise $R_{YY}(s,t)$ would be identically equal to zero. Clearly the solution is unique.

◇ ◇ ◇