Solutions to Homework #6, Math 116

Keziah Cook and Michael McElroy

November 30, 2003

Problem 5.1 (Luenberger)

Define the linear functional f on $L_2[0,1]$ by

$$f(x) = \int_0^1 a(t) \int_0^t b(s)x(s)\mathrm{d}s\mathrm{d}t$$

where $a, b \in L_2[0,1]$. Show that f is a bounded linear functional on $L_2[0,1]$ and find an element $y \in L_2$ such that $f(x) = (x \mid y)$.

Solution to 5.1

- Linearity: $\forall \alpha, \beta \in R, \forall x, y \in L_2[0,1]$, we have

$$\begin{aligned} f(\alpha x + \beta y) &= \int_0^1 a(t) \int_0^t b(s)\left(\alpha x(s) + \beta y(s)\right)\mathrm{d}s\mathrm{d}t \\ &= \int_0^1 a(t)\left(\alpha \int_0^1 \int_0^t b(s)x(s)\mathrm{d}s + \beta \int_0^t b(s)y(s)\mathrm{d}s\right)\mathrm{d}t \\ &= \alpha \int_0^1 \int_0^t b(s)x(s)\mathrm{d}s\mathrm{d}t + \beta \int_0^1 a(t) \int_0^t b(s)y(s)\mathrm{d}s\mathrm{d}t \\ &= \alpha f(x) + \beta f(y) \end{aligned}$$

- Boundedness:

  Lemma. $g(t) = \int_0^t b(s)x(s)\mathrm{d}s \in L_2$

  Proof.

$$\begin{aligned} \left|\int_0^t b(s)x(s)\mathrm{d}s\right| &\leq \int_0^t |b(s)| \cdot |x(s)|\mathrm{d}s \\ &\leq \int_0^1 |b(s)| \cdot |x(s)|\mathrm{d}s \\ &= (|b(s)| \mid |x(s)|) \\ &= k, \text{since } |b|, |x| \in L_2 \end{aligned}$$

  Since $|\int_0^t b(s)x(s)\mathrm{d}s| \leq k$ we know $\|g(t)\| \leq k^2$. Thus $g(t) \in L_2$ as desired. ∎

We can write $f(x) = \int_0^1 a(t)g(t)\mathrm{d}t$. Then

$$\begin{aligned}
|f(x)| &= |\,(a(t) \mid g(t))\,| \\
&= |\left(a(t) \mid \int_0^t b(s)x(s)\mathrm{d}s\right)| \\
&\leq \|a(t)\|_2 \cdot \|\int_0^t b(s)x(s)\mathrm{d}s\|_2, \text{by the Holder inequality} \\
&\leq \|a(t)\|_2 \cdot \|\int_0^t |b(s)||x(s)|\mathrm{d}s\|_2 \\
&\leq \|a(t)\|_2 \cdot \|\int_0^1 |b(s)||x(s)|\mathrm{d}s\|_2 \\
&= \|a(t)\|_2 \cdot \int_0^1 \int_0^1 |b(s)||x(s)|\mathrm{d}s\mathrm{d}t \\
&\leq \|a(t)\|_2 \cdot (|b(s)| \mid |x(s)|) \\
&\leq \|a(t)\|_2 \cdot \|b(s)\|_2 \cdot \|x(s)\|_2 \\
&= k \cdot \|x\|_2
\end{aligned}$$

Thus f is bounded

- Find y such that $f(x) = (x \mid y)$

  To find $y \in L_2$ we apply Fubini's theorem.

$$\begin{aligned}
f(x) &= \int_0^1 a(t) \int_0^t b(s)x(s)\mathrm{d}s\mathrm{d}t \\
&= \int_0^1 \int_0^t a(t)b(s)x(s)\mathrm{d}s\mathrm{d}t \\
&= \int_0^1 \int_s^1 b(s)x(s)a(t)\mathrm{d}t\mathrm{d}s \\
&= \int_0^1 b(s)x(s) \int_s^1 a(t)\mathrm{d}t\mathrm{d}s \\
&= \left(x(s) \mid b(s) \int_s^1 a(t)\mathrm{d}t\right)
\end{aligned}$$

$b(s)\int_s^1 a(t)\mathrm{d}t \in L_2$ because $\int_s^1 a(t)\mathrm{d}t \leq \int_0^1 |a(t)|^2 = k$ for all $s \in [0,1]$. Thus $\|b(s)\int_s^1 a(t)\mathrm{d}t\|_2 < \|b\|_2 \cdot k < \infty$, and $y(s) = b(s)\int_s^1 a(t)\mathrm{d}t$ is the desired element of L_2.

$\square$

Problem 5.2 (Luenberger)

Define the Banach space c as the space of all sequences $x = \{\xi_1, \xi_2, \dots\}$ which converge to a limit, with $\|x\| = sup_{1 \leq k < \infty}|\xi_k|$. Define c_0 as the space of all sequences which converge to zero (same norm as in c). Characterize the dual spaces of c_0 and c (with proofs). Warning: the dual spaces of c_0 and c are not identical.

Solution to 5.2

Claim 1. The dual space of c_0 is (isomorphic to) l_1.

Proof. Let f be a continuous, bounded functional on c_0. We will show that there is a natural way to represent f as an element of l_1. Let $x = \{\xi_1, \xi_2, \dots\} \in c_0$. Define $e_i = \{0, 0, \dots, 1, 0, \dots\}$, where 1 is in the *ith* spot. Let $\eta_i = f(e_i)$. Then by the linearity and continuity of f we can write:

$$\begin{aligned} f(x) &= f\left(\sum_{i=1}^{\infty} e_i \xi_i\right) \\ &= \sum_{i=0}^{\infty} \xi_i \left(f(e_i)\right) \\ &= \sum_{i=0}^{\infty} \xi_i \eta_i \end{aligned}$$

Now define $x_N = \{\xi_1, \xi_2, \dots, \xi_N, 0, ...\}$ where $\xi_i = sgn(\eta_i)$. Recall for $x \in R$,

$$sgn(x) = \begin{cases} +1 & \text{if } x > 0; \\ -1 & \text{if } x < 0; \\ 0 & \text{if } x = 0. \end{cases}$$

Then $f(x_N) = \sum_{i=1}^{N} \mid \eta_i \mid$. But $\mid f(x_N) \mid \leq \|f\| \cdot \|x_N\|$ because f is bounded for all $x \in c_0$. This implies $\mid \sum_{i=1}^{N} |\eta_i| \mid = \sum_{i=1}^{N} |\eta_i| \leq k \cdot (1)$, where $k = \|f\| < \infty$. This holds for all N. Thus $\{\eta_i\} \in l_1$, as desired.

Now we need to show that every $\{\eta_i\} \in l_1$ gives us a bounded linear functional on c_0.

Define $f_{\eta_1}(x)$ for $x \in c_0$ by $\sum_{i=1}^{\infty} \eta_i \xi_i$. Linearity follows by the rules of formal sums. To show boundedness (and thus continuity) we note that

$$\begin{aligned} \mid f_{\eta_i}(x) \mid &= \mid \sum_{i=1}^{\infty} \eta_i \xi_i \mid \\ &\leq |\sum_{i=1}^{\infty} \eta_i (sup_i \xi_i)| \\ &= \|x\| |\sum_{i=1}^{\infty} \eta_i| \\ &= k\|x\|, k < \infty \text{since} \{\eta_i\} \in l_1 \end{aligned}$$

Thus f_{η_1} is bounded.

We showed $c_0^* \subset l_1$ and $l_1 \subset c_0^*$ where '$\subset$' is used up in the isometrically isomorphic sense. □

Claim 2. $c^* \cong l_1 \times R$

Proof. First we note that we can write and $x = \{\xi_1, \xi_1, \dots\} \in c$ where $\xi_i \to r$ as $x = y + \hat{r}$ for $y \in c_0$ and $\hat{r} = \{r, r, r, \dots\}$. In other words, $c = c_0 \times R$. If f is a bounded linear on c, its restriction to $c \times \{0\} = c_0$ can be represented as an element of l_1 by the previous claim. We call this restriction f_1. The restriction of f to $\{0, 0, 0, \dots\} \times R$ is an element of R, since R is self-dual. We call this restriction f_2. We can write

$$\begin{aligned} f(x) &= f(y + \hat{r}) \\ &= f(y) + f(\hat{r}) \\ &= f_1(x) + f_2(x) \end{aligned}$$

Thus we see any f can be represented as a pair (f_1, f_2) where f_1 can be represented as an element of l_1 and f_2 can be represented as an element of R. Thus $c^\star \subset l_1 \times R$ where we use '$\subset$' as before.

Now lets take an arbitrary element $(x, r) \in l_1 \times R$. Define $f_{(x,r)}(p)$ for $p \in c$ where $p = y + \hat{s}$, $y = \{y_1, y_2, \dots\}$ by $f_{(x,r)}(p) = \sum_{i=1}^{\infty} x_i y_i + r\dot{s}$. This is well-defined since $\sum_{i=1}^{\infty} x_i y_i < \infty$ and $r\dot{s} < \infty$. Linearity again follows from properties of formal sums.

To show boundedness, note that

$$\begin{aligned}
\mid f_{(x,r)}(p) \mid &= \mid \sum_{i=1}^{\infty} x_i y_i + r\dot{s} \mid \\
&\leq \mid \sum_{i=1}^{\infty} x_i \sup_i(|y_i|) + r\dot{s} \mid \\
&\leq \mid \sum_{i=1}^{\infty} x_i \sup_i(|p_i|) + r \cdot \sup_i(|p_i|) \mid \\
&= \mid \sup_i |p_i| \left(\sum_{i=1}^{\infty} x_i + s \right) \mid \\
&\leq \|p\| \left(\|x\|_1 + s \right) \\
&= \|p\| \cdot k, \quad k < \infty
\end{aligned}$$

Thus $f_{(x,r)}$ is bounded, and any element in $l_i \times R$ represents a bounded linear functional on c. Thus $l_1 \times R \subset c^\star$ and $c^\star$ is isomorphic to $l_1 \times R$. □

Note $l_1 \times R$ is isomorphic to l_1 (just prepend r to $\{\xi_1, \xi_2, \dots\}$. However the map from l_1 to $c^\star$ is different from the map from l_1 to $c_0^\star$. So the duals to c and c_0 are different.

□

Problem Writing 1 (Goroff)

Consider table 1.

Year	Employees	Applications	Consultations
1	93	5804	1750
2	100	6250	1878
3	102	6200	1852
4	105	6225	1860

We want to model the number of employees from the data of work accomplished: applications processed and consultations provided. We want to fit the data in table 1 to a linear model $f(\xi_1, \xi_2) = b_0 + b_1\xi_1 + b_2\xi_2$.

(a) Describe how to determine the $\{b_i\}$ to best fit the data in a least-squares sense.

(b) Compute the solution.

Solution to Writing 1

(a) First, let's define some variables. Let $y_0 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$, $y_1 = \begin{pmatrix}5804\\6250\\6200\\6225\end{pmatrix}$, $y_2 = \begin{pmatrix}1750\\1878\\1852\\1860\end{pmatrix}$, and $p = \begin{pmatrix}93\\100\\102\\105\end{pmatrix}$.

Also define the vector b of parameters to be estimated: $b = \begin{pmatrix}b_0\\b_1\\b_2\end{pmatrix}$. Define the matrix A as having the y_i's for columns: $A = [y_0, y_1, y_2]$. We want a model to minimize $\|p - b_0y_0 - b_1y_1 - b_2y_2\|^2 = (p - Ab)^T(p - Ab)$. We are projecting p, a 4-vector, onto the 3-dimensional subspace spanned by the columns of A. Theorem 1 on page 83 of Luenberger suggests the formula $b^* = (A^TA)^{-1}A^Tp$.

(b)

$$b^* = \begin{pmatrix}b_0\\b_1\\b_2\end{pmatrix} \approx \begin{pmatrix}27.68\\0.1153\\-0.3452\end{pmatrix}$$

$$= \left[\begin{pmatrix}1 & 1 & 1 & 1\\5804 & 6250 & 6200 & 6225\\1750 & 1878 & 1852 & 1860\end{pmatrix}\begin{pmatrix}1 & 5804 & 1750\\1 & 6250 & 1878\\1 & 6200 & 1852\\1 & 6225 & 1860\end{pmatrix}\right]^{-1}\begin{pmatrix}1 & 1 & 1 & 1\\5804 & 6250 & 6200 & 6225\\1750 & 1878 & 1852 & 1860\end{pmatrix}\begin{pmatrix}93\\100\\102\\105\end{pmatrix}.$$

□

Problem Writing 2 (Goroff)

Let $A \in \mathbb{R}^{n\times n}$, and consider the inner product on the space of $n \times n$ matrices $\langle A|B\rangle = \operatorname{trace}(A^TB)$. Find the constant c such that cI best approximates A in the least-squares sense.

Solution to Writing 2

Let $M = \{\alpha I,\ \alpha \in \mathbb{R}\}$ be the subspace of scaled identity matrices, in which cI dwells. So we are projecting A onto M. Furthermore, M is obviously closed, so our projection theorem holds.

Idea 1 Projection theorem. We have that if c is optimal then $\forall\alpha \in \mathbb{R}$, $\langle A - cI|\alpha I\rangle = 0$.

$$\alpha\,(\operatorname{trace}(A) - cn) = 0 \implies c = \operatorname{trace}(A)/n.$$

Idea 2 Calculus. Let $f(c) = \langle A - cI|A - cI\rangle$. Then $f(c) = \operatorname{trace}(A^TA) - 2c\cdot\operatorname{trace}(A) + c^2n$.

$$0 = \mathrm{d}f/\mathrm{d}c = -2\cdot\operatorname{trace}(A) + 2cn \implies c = \operatorname{trace}(A)/n \text{ is a critical point.}$$

$\frac{\mathrm{d}^2f}{\mathrm{d}c^2}\Big|_{c=\operatorname{tr}A/n} = 2n$, so this is the (unique, global) minimum point.

□

Problem Writing 3 (Goroff)

Find the least-squares solution for the system

$$\begin{pmatrix}1 & 2\\2 & 1\\4 & 5\end{pmatrix}\begin{pmatrix}\xi_1\\\xi_2\end{pmatrix} \approx \begin{pmatrix}3\\3\\10\end{pmatrix}.$$

Solution to Writing 3

(a) Gram-Schmidt. We are projecting the vector $p = [3, 3, 10]^T$ onto the plane spanned by the columns of $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 4 & 5 \end{pmatrix}$. We need an orthogonal basis for this subspace: if $y_0 = [1, 2, 4]^T$ and $y_1 = [2, 1, 5]^T$ then choose $z_0 = y_0$ and $z_1 = y_1 - \frac{y_1^T z_0}{z_0^T z_0} z_0 = y_1 - \frac{8}{7} y_0$. Now project p onto z_0 and z_1: $\hat{p} = \alpha_0 z_0 + \alpha_1 z_1$, $\alpha_i = \frac{p^T z_i}{z_i^T z_i}$. That is, $\alpha_0 = 49/21 = 7/3$ and $\alpha_2 = 7/6$. Finally, translate these coordinates in terms of z_i into coordinates for $\hat{p}$ in terms of y_0, y_1.

$$\hat{p} = \begin{pmatrix} \alpha_0 & \alpha_1 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} & \frac{7}{6} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{8}{7} & 1 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & \frac{7}{6} \end{pmatrix}}_{=(\xi_1\ \xi_2)} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}.$$

(b) Normal equations. The normal equations are succiently written in matrix form as

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = (A^T A)^{-1} A^T p = \begin{pmatrix} 1 \\ \frac{7}{6} \end{pmatrix}.$$

□

Problem Writing 4 (Goroff)

We know a few statistics for a sequence of random variables x_i, namely that the expected values of $x_n, x_n^2, x_n x_{n-1}$, and $x_n x_{n-2}$ are respectively $1, \frac{9}{4}, \frac{7}{4}$, and $\frac{5}{4}$. We want the coefficients minimizing the expected squared error of a linear predictor function based the past two measurements. That is, choose (ξ_0, ξ_1, ξ_2) such that $\mathbf{E}(x_n - \hat{x}_n)^2$ is minimized; $\hat{x}_n = \begin{pmatrix} \xi_0 & \xi_1 & \xi_2 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ x_{n-1} \\ x_{n-2} \end{pmatrix}$.

Solution to Writing 4

Idea 1 Gram-Schmidt. We want to project x_n onto a set of orthogonal basis vectors for the subspace $\text{span}\{\mathbf{1}, x_{n-1}, x_{n-2}\}$. Again, normalizing your basis vectors is a waste of time! First, create the basis $\{z_i\}$:

$$\begin{aligned} z_0 &= \mathbf{1}, \\ z_1 &= x_{n-1} - \frac{\langle x_{n-1} | z_0 \rangle}{\langle z_0 | z_0 \rangle} z_0 = x_{n-1} - \mathbf{1}, \\ z_2 &= x_{n-2} - \frac{\langle x_{n-2} | z_1 \rangle}{\langle z_1 | z_1 \rangle} z_1 - \frac{\langle x_{n-2} | z_0 \rangle}{\langle z_0 | z_0 \rangle} z_0 = x_{n-2} - \frac{\mathbf{E}(x_{n-1} x_{n-2} - x_{n-2})}{\mathbf{E}(x_{n-1}^2 - 2x_{n-1} + \mathbf{1})} (x_{n-1} - \mathbf{1}) - \mathbf{1} \\ &= x_{n-2} - \frac{3}{5} x_{n-1} - \frac{2}{5} \mathbf{1}. \end{aligned}$$

Second, project x_n onto this basis: $\hat{x}_n = \sum \alpha_i z_i$, $\alpha_i = \frac{\langle x_n | z_i \rangle}{\langle z_i | z_i \rangle} = \frac{\mathbf{E} x_n z_i}{\mathbf{E} z_n^2}$. Hence $(\alpha_0, \alpha_1, \alpha_2) = (1, \frac{3}{5}, -\frac{1}{4})$. Lastly, we change coordinates back to using $(\mathbf{1}, x_{n-1}, x_{n-2})$ as a basis:

$$\hat{x}_n = \begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{3}{5} & -\frac{1}{4} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -\frac{2}{5} & -\frac{3}{5} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{1} \\ x_{n-1} \\ x_{n-2} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{2} & \frac{3}{4} & -\frac{1}{4} \end{pmatrix}}_{=(\xi_1\ \xi_2\ \xi_3)} \begin{pmatrix} \mathbf{1} \\ x_{n-1} \\ x_{n-2} \end{pmatrix}.$$

Idea 2 Normal equations. Previously we used the trick[1] that $\xi^* = (A^T A)^{-1} A^T p$, but will that work here? Our inner product is the expected value and not just the vector inner product. However, if we remember to take expectations to make sense of our random variables, then we are fine. $\hat{x}_n = \begin{pmatrix} \mathbf{1} & x_{n-1} & x_{n-2} \end{pmatrix} \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \end{pmatrix}$. Let $p = x_n$, $A = \begin{pmatrix} \mathbf{1} & x_{n-1} & x_{n-2} \end{pmatrix}$. Then we have

$$
\begin{aligned}
\xi^* = \begin{pmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \end{pmatrix} &= \left(\mathbf{E} A^T A\right)^{-1} \mathbf{E}(A^T p) \\
&= \left[\mathbf{E}\begin{pmatrix} \mathbf{1} & x_{n-1} & x_{n-2} \\ x_{n-1} & x_{n-1}^2 & x_{n-1}x_{n-2} \\ x_{n-2} & x_{n-1}x_{n-2} & x_{n-2}^2 \end{pmatrix}\right]^{-1} \mathbf{E}\begin{pmatrix} x_n \\ x_{n-1}x_n \\ x_{n-2}x_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{9}{4} & \frac{7}{4} \\ 1 & \frac{7}{4} & \frac{9}{4} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ \frac{7}{4} \\ \frac{5}{4} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{3}{4} \\ -\frac{1}{4} \end{pmatrix}.
\end{aligned}
$$

Idea 3 Projection theorem. We are trying to project x_n into the subspace spanned by $\mathbf{1}$, and the previous two measurements. If we have the optimal ξ_i, then the difference will be orthogonal to the subspace, in particular, $x_n \hat{x}_n \perp \mathbf{1}, x_{n-1}, x_{n-2}$. that is,

$$0 = \langle x_n - \xi_0 \mathbf{1} - \xi_1 x_{n-1} - \xi_2 x_{n-2} | \mathbf{1}, x_{n-1}, x_{n-2} \rangle.$$

Expanding this out gives us our normal equations above.

□

Problem Writing 5 (Goroff)

Find the continuous function x for which $\int_0^1 x(t)\mathrm{d}t = 1$ and $\int_0^1 x^2(t)\mathrm{d}t$ is minimum.

Solution to Writing 5

Consider $H = L_2[0,1]$ with $(x \mid y) = \int_0^1 x(t)y(t)\mathrm{d}t$ as the inner product. We know H is a Hilbert space. The problem we want to solve is:

$$\min_{x \in H} \|x\|_2 \text{s.t}(x|\mathbf{1}) = 1, \text{where } \mathbf{1}(t) = 1 \forall t \tag{1}$$

Theorem 2 in Luenberger (pg. 65) tells us that the solution to 1 satisfies

1. $x_0 = \beta \mathbf{1}$ for some $\beta \in R$ and
2. $(x_0 \mid \mathbf{1}) = 1$

Plugging in, we get

$$
\begin{aligned}
(x_0 \mid \mathbf{1}) &= (\beta \mathbf{1} \mid \mathbf{1}) && (2) \\
&= \beta(\mathbf{1} \mid \mathbf{1}) && (3) \\
&= \beta && (4)
\end{aligned}
$$

Thus, $\beta = 1$, and $x(t) = \mathbf{1}$ is the solution to the minimization problem. □

Problem Writing 6 (Goroff)

[1] Theorem 1 on page 83 of Luenberger

Define a new inner product on $L_2\,[0,1]$ by setting $(x|y) = \int_0^1 e^t x(t)y(t)\mathrm{d}t$ Using this weighted inner product to define a norm, reconsider Example 1 on page 66 of Luenberger and find the control voltage function u that solves the motor moving problem given there but minimizes this norm. Explain what effect this change has on the previous energy minimizing solution.

Solution to Writing 6

The original problem states that the motor begins at rest. So $\theta(0) = \omega(0) = 0$. We want to find the minimum energy function (with respect to our new norm) which rotates the shaft to $\theta(1) = 1$ and $\omega(1) = 0$. As on Luenberger page 66, this gives us two constraints that $u(t)$ must satisfy:

$$\begin{aligned}\omega(1) &= \int_0^1 e^{(t-1)}u(t)\mathrm{d}t = 0 \\ \theta(1) &= \int_0^1 \{1 - e^{(t-1)}\}u(t)\mathrm{d}t = 1\end{aligned}$$

Since we want to treat u as an element of L_2 with our new norm, we need to find y_1, y_2 such that the above restrictions can be expressed as:

$$\begin{aligned}\omega(1) &= (y_1 \mid u) \\ \theta(1) &= (y_2 \mid u)\end{aligned}$$

We re-write our two constraints as:

$$\begin{aligned}\omega(1) &= \int_0^1 e^t \cdot e^{-1}u(t)\mathrm{d}t = 0 \\ \theta(1) &= \int_0^1 e^t\{e^{-t} - e^{-1}\}u(t)\mathrm{d}t = 1\end{aligned}$$

It is easy to see that with respect to our new inner product, $y_1 = e^{-1}$ and $y_2 = \{e^{-t} - e^{-1}\}$.

Theorem 2 (page 65) tells us that $u(t) = \alpha y_1(t) + \beta y_2(t)$, where α, β are such that the constraints are satisfied. Plugging $u(t)$ into the first constraint yields:

$$\begin{aligned}\omega(1) &= \int_0^1 e^t \cdot e^{-1}\{\alpha y_1(t) + \beta y_2(t)\}\mathrm{d}t \\ &= \int_0^1 e^{t-1}\{\alpha e^{-1} + \beta(e^{-t} - e^{-1})\}\mathrm{d}t \\ &= \int_0^1 (\alpha - \beta)e^{t-2}\mathrm{d}t + \beta\int_0^1 e^{-1}\mathrm{d}t \\ &= (\alpha - \beta)(e^{-1} - e^{-2}) + \beta e^{-1} \\ &= \alpha(e^{-1} - e^{-2}) + \beta e^{-2} = 0 \\ \Rightarrow \beta &= \alpha(1 - e)\end{aligned}$$

Plugging into the second constraint yields:

$$\begin{aligned}\theta(1) &= \int_0^1 \{1 - e^{t-1}\}\{\alpha y_1(t) + \beta y_2(t)\}\mathrm{d}t \\ &= \int_0^1 (\alpha + 2\beta)e^{-1}\mathrm{d}t + \int_0^1 (-2\beta)e^{-t}\mathrm{d}t - \int_0^1 \alpha e^{t-2}\mathrm{d}t \\ &= (\alpha + 2\beta)e^{-1} + 2\beta(e^{-1} - 1) - \alpha(e^{-1} - e^{-2}) = 1\end{aligned}$$

We plug in $\beta = \alpha(1-e)$ and get:

$$\begin{aligned}
(\alpha + 2\alpha(1-e))e^{-1} + 2\alpha(1-e)(e^{-1}-1) - \alpha(e^{-1} - e^{-2}) &= 1 \\
e^{-2}\left(4\alpha e - 4\alpha e^2 + 2\alpha e^3 + \alpha\right) &= 1 \\
\frac{e^2}{2e^3 - 4e^2 + 4e + 1} &= \alpha \\
\frac{e^2(1-e)}{2e^3 - 4e^2 + 4e + 1} &= \beta \\
\frac{e}{2e^3 - 4e^2 + 4e + 1} + \frac{e^2(1-e)}{2e^3 - 4e^2 + 4e + 1}(e^{-t} - e^{-1}) &= u(t)
\end{aligned}$$

Simplifying yields:

$$\begin{aligned}
u(t) &= \frac{e^2 - e^3}{2e^3 - 4e^2 + 4e + 1}e^{-t} + \frac{e^2}{2e^3 - 4e^2 + 4e + 1} \\
&\approx -.5646e^{-t} + .3286
\end{aligned}$$

Some students noticed that $[e^{-1}, e^{-t} - e^{-1}] = [1, e^{-t}]$ and plugged $u(t) = \alpha + \beta e^{-t}$ into the constraints. This led to a slightly more straightforward calculation of $u(t)$.

To understand why and how this solution differs from the original solution to Example 2 in Luenberger, we just need to think about how the new inner product "weights" $u(t)$. e^t increases from 1 to $e \approx 2.718$ on $[0,1]$. Thus is it more costly to apply thrust near $t=1$ than it is near $t=0$. So our optimal thrust function $u(t)$ has thrust decreasing from $t=0$ to $t=1$ as e^{-t} decays.

□