Solutions to Homework #8, Math 116

Keziah Cook and Michael McElroy

January 20, 2004

1 Theorems

Theorem (Minimum norm duality theorem). *Let $x \in X$ be given, where X is a normed linear space, and let d denote its distance from a subspace M. Then*

$$d = \inf_{m \in M} \|x - m\| = \max_{\substack{\|x^*\| \le 1 \\ x^* \in M^\perp}} \langle x, x^* \rangle$$

where the maximum is achieved for some $x_0^* \in M^\perp$. If the infimum is achieved for some $m_0 \in M$, then x_0^* is aligned with $x - m_0$, i.e. $\langle x - m_0, x_0^* \rangle = \|x - m_0\|_X \cdot \|x_0^*\|_{X^*}$.

Theorem (Fenchel duality theorem). *Let f and g be, respectively, convex and concave functionals on convex sets C and D in a normed space X. Assume that $C \cap D$ has points in the relative interior of C and D and that either $[f, C]$ or $[g, D]$ has non-empty interior. Then if $\inf_{x \in C \cap D} \{f(x) - g(x)\}$ is finite we have*

$$\mu = \inf_{x \in C \cap D} \{f(x) - g(x)\} = \max_{x^* \in C^* \cap D^*} \{g^*(x^*) - f^*(x^*)\}$$

where the maximum on the right is acheived by some $x_0^* \in C^* \cap D^*$.

Furthermore, if the infimum on the left is achieved for some $x_0 \in C \cap D$, then x_0 maximizes $[\langle x, x_0^* \rangle - f(x)]$ for $x \in C$ and x_0 minimizes $[\langle x, x_0^* \rangle - g(x)]$ for $x \in D$.

Theorem (Lagrange duality theorem). *Let X be a vector space, Z be a normed space, and $\Omega \subset X$ be convex. Let $f : \Omega \to \mathbb{R}$ and $G : X \to Z$ be convex operators. Suppose there exists $x_1 \in X$ such that $G(x_1) < 0$ and that $\inf \{ f(x) \mid G(x) \le 0, x \in \Omega \} < \infty$. Then*

$$\inf_{\substack{G(x) \le 0 \\ x \in \Omega}} f(x) = \max_{z^* \ge 0} \phi(z^*)$$

and the maximum on the right is achieved by some $z_0^* \ge 0$.

If the infimum on the left is achieved for some $x_0 \in \Omega$, then $\langle G(x_0), z_0^* \rangle = 0$ and x_0 minimizes $f(x) + \langle G(x), z_0^* \rangle$, $x \in \Omega$.

2 Problems and solutions

Problem A and B (Goroff)

Find a curve x joining the given endpoints that extremizes J:

A. $J[x] = \int_0^2 \frac{\dot{x}^2}{x^3}\,\mathrm{d}t$, with $x(0) = 1$ and $x(2) = 4$. There may seem to be two solutions. How can you rule one out?

B. $J[x] = \int_0^1 \frac{\sqrt{1+\dot{x}^2}}{x}\,\mathrm{d}t$ with $x(0) = 0$ and $x(1) = \sqrt{3}$.

Solution to A and B

Remember the Euler-Lagrange equations: to extremize $\int_0^T F(t, x, \dot{x})\,\mathrm{d}t$ we require that $\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{x}} = 0$. If $F = F(x, \dot{x})$ then the E-L equation integrates to $F - \dot{x}\frac{\partial F}{\partial \dot{x}} = c$. It is this latter equation that we use to solve A and B.

Here, the Euler-Lagrange equations will leave us with some equation $g(x, \dot{x}) = 0$. We wish to solve for $\dot{x}$, say, $\dot{x} = h(x)$. Then separation of variables and an integration leaves us with an equation of x and t: $\frac{\mathrm{d}x}{h(x)} = \mathrm{d}t$ so $\int \frac{1}{h(x)}\,\mathrm{d}x = t + c$.

A. $F = \frac{\dot{x}^2}{x^3}$, so

$$F - \dot{x}\frac{\partial F}{\partial \dot{x}} = c_1 \quad \Longrightarrow \quad \frac{\dot{x}^2}{x^3} - \dot{x} \cdot \frac{2\dot{x}}{x^3} = -\frac{\dot{x}^2}{x^3} = c_1 \quad \Longrightarrow \quad \dot{x}^2 = -c_1 x^3.$$

Taking square roots and rearranging yields

$$x^{-3/2}\mathrm{d}x = c_2\mathrm{d}t.$$

Integrating yields

$$-\frac{2}{\sqrt{x}} = c_2 t + d_1 \implies x(t) = \frac{1}{(ct + d)^2}. \tag{1}$$

Our boundary conditions leads to $d = 1$ and $c = -\frac{1}{4}$ or $-\frac{3}{4}$. However, if $c = -\frac{3}{4}$, then $ct + d = 0$ for $t = \frac{4}{3}$, a singularity where $x(t)$ would blow up. Hence we choose $c = -\frac{1}{4}$, and

$$x(t) = \frac{1}{\left(-\frac{t}{4} + 1\right)^2} = \frac{16}{(4 - t)^2}.$$

A'. Now we'll solve problem A using the original Euler-Langrange equation, $\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{x}} = 0$. This will require an additional integration than if we'd use $F - \dot{x}\frac{\partial F}{\partial \dot{x}} = c$, but it will show some additional integration tricks which you may find useful. $F = \dot{x}^2 x^{-3}$, so

$$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial F}{\partial \dot{x}} = 0 \quad \Longrightarrow \quad -3\dot{x}^2 x^{-4} - \frac{\mathrm{d}}{\mathrm{d}t}\left(2\dot{x}x^{-3}\right) = 0 \quad \Longrightarrow \quad -3\dot{x}^2 x^{-4} - 2\ddot{x}x^{-3} + 6\dot{x}^2 x^{-4} = 0$$

$$\Longrightarrow 3\dot{x}^2 x^{-4} - 2\ddot{x}x^{-3} = 0 \quad \Longrightarrow \quad 3\left(\frac{\dot{x}}{x}\right)^2 - 2\left(\frac{\ddot{x}}{x}\right) = 0.$$

Now use the substitution $y = \frac{\dot{x}}{x}$. We have $\dot{y} = \frac{\ddot{x}}{x} - \frac{\dot{x}^2}{x^2}$. So $\frac{\ddot{x}}{x} = \dot{y} + y^2$. Substituting this into our last equation,

$$3\left(\frac{\dot{x}}{x}\right)^2 - 2\left(\frac{\ddot{x}}{x}\right) = 0 \quad \Longrightarrow \quad 3y^2 - 2\left(\dot{y} + y^2\right) = 0 \quad \Longrightarrow \quad y^2 - 2\dot{y} = 0$$

$$\Longrightarrow 2\dot{y}y^{-2} = 1 \quad \Longrightarrow \quad 2y^{-2}\mathrm{d}y = \mathrm{d}t$$

and integration yields

$$-2y^{-1} = t + c \quad \text{or} \quad y(t) = -\frac{2}{t + c}.$$

Now replacing $y(t)$ with $\dot{x}x^{-1}$ we have

$$\dot{x}x^{-1} = y(t) = -\frac{2}{t+c} \quad \Longrightarrow \quad \frac{\mathrm{d}x}{x} = -\frac{2}{t+c}\mathrm{d}t$$

which when integrated yields

$$\log x(t) = -2\log(t+c) + d', \quad \text{or} \quad x(t) = \frac{d}{(t+c)^2}. \tag{2}$$

Note that (2) has the same form as (1), our solution in **A**.

B. $F = \frac{\sqrt{1+\dot{x}^2}}{x}$, so

$$F - \dot{x}\frac{\partial F}{\partial \dot{x}} = c' \implies \frac{\sqrt{1+\dot{x}^2}}{x} - \dot{x} \cdot \frac{2\dot{x}}{2x\sqrt{1+\dot{x}^2}} = c'.$$

Multiplying both sides by $x\sqrt{1+\dot{x}^2}$ yields

$$1 = c'x\sqrt{1+\dot{x}^2} \implies 1 = c'^2x^2\left(1+\dot{x}^2\right)$$

$$\implies \dot{x}^2 = \frac{c^2-x^2}{x^2} \implies \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\sqrt{c^2+x^2}}{x} \implies \frac{x}{\sqrt{c^2+x^2}}\mathrm{d}x = \mathrm{d}t$$

Integrating the last equation is tricky, but here we apply the same trig substitution as with the brachistochrone problem, namely $x = c\sin\theta$, $\mathrm{d}x = c\cos\theta\,\mathrm{d}\theta$. Now integrating is a snap:

$$t + d = \int \mathrm{d}t = \int \frac{x}{\sqrt{c^2-x^2}}\mathrm{d}x = \int \frac{c\sin\theta}{c\cos\theta} \cdot c\cos\theta\,\mathrm{d}\theta = -c\cos\theta.$$

Since $\sin^2 + \cos^2 = 1$, we have

$$x^2 + (t+d)^2 = c^2\sin^2\theta + c^2\cos^2\theta = c^2.$$

Our extremal curve is the arc of a circle. $x(0) = 0$ implies $d^2 = c^2$, while $x(1) = \sqrt{3}$ implies $d = -2$. Hence our equation is $x^2(t) + (t-2)^2 = 4$, or

$$x(t) = \sqrt{4-(t-2)^2} = \sqrt{4t-t^2}.$$

□

Problem 7.13 (Luenberger)

Candidates allocation problem: two candidates X and Y for political office wish to distribute their advertizing dollars in n markets so as to gain the most votes. Assume that their allocations are bounded above: $\sum_{i=1}^n x_i = x_0$, $\sum_{i=1}^n y_i = y_0$, and that the relationship between spending x_i, y_i in a market with u_i undecided votes is to gain $\frac{x_i-y_i}{x_i+y_i}u_i$ votes (votes are tabulated positive for X, negative for Y. Find their strategies to optimize their output.

Solution to 7.13

Before we solve the problem, let me make a few statements about what our intuition tells us about the solution. The problem is symmteric with respect to X and Y, so we expect that their strategies are identical (we will prove this, of course). Furthermore, let us produce an optimal vote function v,

$$v(x_0, y_0) = \sum_{i=1}^{n} \frac{x_i - y_i}{x_i + y_i}u_i.$$

Obviously, symmtery yields $v(x_0, y_0) = -v(y_0, x_0)$, so the vote function is skew-symmetric with respect to its arguments. Furthermore, intuitively we have $\operatorname{sign}(x_0 - y_0) = \operatorname{sign}(v(x_0, y_0))$ because whoever has the most dollars can choose to outspend the other in every market.

We use Lagrange multipliers when solving a problem of the form $\max f(x)$ such that $g(x) = c$. In this case, define the Lagrangian

$$L(x, y, \lambda) = \sum_{i=1}^{n} \frac{x_i - y_i}{x_i + y_i} u_i + \begin{pmatrix} x_0 - \sum_{i=1}^n x_i & y_0 - \sum_{i=1}^n y_i \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}.$$

Now we have $\frac{\partial L}{\partial x_i} = \frac{\partial L}{\partial y_i} = 0$ and $\frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = 0$. The first two derivatives equalling 0 optimize the function while the latter two enforce the two constraints.

$$0 = \frac{\partial L}{\partial x_i} \implies \frac{2y_i u_i}{(x_i + y_i)^2} = \lambda_1, \tag{3}$$

and

$$0 = \frac{\partial L}{\partial y_i} \implies \frac{2x_i u_i}{(x_i + y_i)^2} = \lambda_2. \tag{4}$$

Combining the results of (3) and (4), we get $\frac{(x_i+y_i)^2}{2u_i} = \frac{y_i}{\lambda_1} = \frac{x_i}{\lambda_2}$, so $y_i \propto x_i$, or

$$y_i = \frac{y_0}{x_0} x_i.$$

Substituting this in for y_i into (3), we get

$$\lambda_1 = \frac{2y_i u_i}{(x_i + y_i)^2} = \frac{2\frac{y_0}{x_0} x_i u_i}{\frac{(x_0+y_0)^2}{x_0^2} x_i^2} = \frac{2x_0 y_0 u_i}{(x_0 + y_0)^2 x_i}.$$

This gives our solution, for

$$\frac{x_i}{u_i} = \frac{2x_0 y_0}{\lambda_1 (x_0 + y_0)^2} = \text{const} \quad \implies \quad x_i \propto u_i.$$

□

Problem 7.14 (Luenberger)

Let $X = L_2[0,1]$ and define $f : L_2[0,1] \to \mathbb{R}$, $f(x) = \int_0^1 \frac{1}{2}x^2(t) + |x(t)| \,\mathrm{d}t$. Find the conjugate functional of f.

Solution to 7.14

Let $C = X = L_2[0,1]$. Since L_2 is a Hilbert space, $X^* = X = L_2[0,1]$, and for $x, x^* \in L_2[0,1]$ we have $\langle x, x^* \rangle = \langle x \mid x^* \rangle = \int_0^1 x(t)x^*(t) \,\mathrm{d}t$. Before we address f^*, we need to know its domain.

$$\begin{aligned} C^* &= \left\{ y \in L_2[0,1] \;\middle|\; \sup_{x \in L_2[0,1]} [\langle x, y \rangle - f(x)] < \infty \right\} \\ &= \left\{ y \in L_2[0,1] \;\middle|\; \sup_{x \in L_2[0,1]} \int_0^1 x(t)y(t) - \frac{1}{2}x^2(t) - |x(t)| \,\mathrm{d}t < \infty \right\}. \end{aligned}$$

Noting that Cauchy-Schwartz-Buniakovsky gives us

$$\int_0^1 x(t)y(t) - \frac{1}{2}x^2(t) - |x(t)|\,\mathrm{d}t \underbrace{\leq}_{CSB} \|x\|_2\|y\|_2 - \frac{1}{2}\|x\|_2^2 - \|x\|_1 \leq \|x\|_2\|y\|_2 - \frac{1}{2}\|x\|_2^2 \leq \frac{1}{2}\|y\|_2^2 < \infty.$$

So $C^* = C = L_2[0,1]$.

The conjugate functional is $f^* : L_2[0,1] \to \mathbb{R}$,

$$f^*(y) = \sup_{x \in L_2[0,1]} \int_0^1 x(t)y(t) - \frac{1}{2}x^2(t) - |x(t)|\,\mathrm{d}t \leq \sup_{x \in L_2[0,1]} \left\{ \|x\|_2\|y\|_2 - \frac{1}{2}\|x\|_2^2 - \|x\|_1 \right\}. \tag{5}$$

To make (5) into an equality, Cauchy-Schwartz tells us to choose $x = ky$, where $k \in \mathbb{R}$, $k \geq 0$. That is, choose x aligned with y:

$$f^*(y) = \sup_{k \geq 0} \left\{ \left(k - \frac{k^2}{2}\right) \cdot \|y\|_2^2 - k \cdot \|y\|_1 \right\}.$$

Calculus yields

$$k = \frac{\|y\|_2^2 - \|y\|_1}{\|y\|_2^2}; \qquad f^*(y) = \frac{\left(\|y\|_2^2 - \|y\|_1\right)^2}{2\|y\|_2^2} = \frac{1}{2}\left(\|y\|_2 - \frac{\|y\|_1}{\|y\|_2}\right)^2.$$

□

Problem 7.17 (Luenberger)

Let f be a convex functional on a convex set C in a normed space and let $[f^*, C^*] = [f, C]^*$. For $x \in C$, $x^* \in C^*$, deduce Young's inequality

$$\langle x, x^* \rangle \leq f(x) + f^*(x^*).$$

Apply this result to L_p spaces.

Solution to 7.17

Let $x \in C$, $x^* \in C^*$ be given. By the definition of f^*, we have

$$f(x) + f^*(x^*) = f(x) + \sup_{y \in C} [\langle y, x^* \rangle - f(x)] \geq f(x) + \langle x, x^* \rangle - f(x) = \langle x, x^* \rangle.$$

This is Young's inequality.

Hölder's inequality for ℓ_p

We will derive Hölder's inequality by judicious choice of f. Choose $f(x) = \frac{1}{p}x^p$. C will be the positive reals $\mathbb{R}^+$. Now

$$C^* = \left\{ y \in \mathbb{R} \;\middle|\; \sup_{x \in \mathbb{R}^+} \left[xy - \frac{1}{p}x^p \right] < \infty \right\} = \mathbb{R}$$

and f^* is defined via

$$f^*(y) = \sup_{x \in \mathbb{R}^+} \left[xy - \frac{1}{p}x^p \right] = \begin{cases} 0 & : \quad y \leq 0 \\ \frac{p-1}{p} y^{p/(p-1)} & : \quad y > 0 \end{cases}.$$

Now we have

$$f(x) + f^*(y) = \frac{1}{p}x^p + \frac{p-1}{p}y^{p/(p-1)} = \frac{1}{p}x^p + \frac{1}{q}y^q,$$

for $\frac{1}{p} + \frac{1}{q} = 1$.

Consider $X = \ell_p$, $1 < p < \infty$. For $a = \{\xi_1, \xi_2, \ldots\} \in \ell_p$ and $b = \{\eta_1, \eta_2, \ldots\} \in \ell_q$, choosing $x = \frac{|\xi_i|}{\|a\|_p}$ and $y = \frac{|\eta_i|}{\|b\|_q}$ gives us for each i

$$\langle x, y \rangle \leq f(x) + f^*(y) \quad \implies \quad \frac{|\xi_i \eta_i|}{\|a\|_p \|b\|_q} \leq \frac{1}{p}\left(\frac{|\xi_i|}{\|a\|_p}\right)^p + \frac{1}{q}\left(\frac{|\eta_i|}{\|b\|_q}\right)^q.$$

Summing this inequality over i yields the Hölder inequality,

$$\frac{\sum_{i=1}^{\infty} |\xi_i \eta_i|}{\|a\|_p \|b\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1, \quad \text{or} \quad |\langle a \mid b \rangle| \leq \|a\|_p \|b\|_q.$$

Hölder's inequality for L_p

It is not hard to show the same result for $X = L_p$, $1 < p < \infty$. Clearly $X^* = L_q$, $1/p + 1/q = 1$. Choose $f(x) = \|x\|_p$; for $x^* \in L_q$ the conjugate functional is

$$f^*(x^*) = \sup_x \left[\langle x, x^* \rangle - \|x\|_p\right] \leq \|x\|_p \left(\|x^*\|_q - 1\right).$$

If $f^*(x^*)$ is to be finite, we must have $\|x^*\|_q \leq 1$, in which case $f^*(x^*) = 0$. $D^* = \left\{x^* \mid \|x^*\| \leq 1\right\}$. Now if $\|x^*\| \leq 1$, Young's inequality yields

$$\langle x, x^* \rangle \leq f(x) + f^*(x^*) = \|x\|_p. \tag{6}$$

What about $x^* \in L_q$, $\|x^*\| > 1$? We may extend (**??**) to all $x^* \in L_q$ via

$$\langle x, x^* \rangle = \left\langle x, \frac{x^*}{\|x^*\|_q} \right\rangle \cdot \|x^*\|_q \leq \|x\|_p \cdot \|x^*\|_q.$$

□

Problem 7.18 (Luenberger)

Use the Fenchel duality theorem to prove the minimum norm duality theorem from section 5.8 of Luenberger.

Solution to 7.18

Let $C = X$, $D = M$, $f(y) = \|x - y\|$, $g(y) = 0$. Then C^* and D^* are

$$C^* = \left\{x^* \in X^* \;\middle|\; \sup_{y \in X}\left[\langle y, x^* \rangle - \|x - y\|\right] < \infty\right\} = \left\{x^* \in X^* \mid \|x^*\| \leq 1\right\},$$

$$D^* = \left\{x^* \in M^* \;\middle|\; \inf_{y \in M}\left[\langle y, x^* \rangle - 0\right] > -\infty\right\} = M^\perp.$$

The functionals conjugate to f and g are

$$f^* : C^* \to \mathbb{R}; \quad f^*(x^*) = \sup_{y \in X}\left[\langle y, x^* \rangle - \|x - y\|\right] = \langle x, x^* \rangle \tag{7}$$

and

$$g^* : M^\perp \to \mathbb{R}; \quad g^*(x^*) = \inf_{m\in M} \langle m, x^* \rangle = 0.$$

Plugging our terms into the Fenchel duality theorem yields

$$\inf_{m\in M} \|x - m\| = \max_{\substack{\|x^*\| \le 1 \\ x^* \in M^\perp}} -\langle x, x^* \rangle = \max_{\substack{\|x^*\| \le 1 \\ x^* \in M^\perp}} \langle x, x^* \rangle, \tag{8}$$

which is the first statement of the minimum norm duality theorem.

To prove the alignment result, choose $x_0^* = \arg \max_{x^* \in C^* \cap D^*} \underbrace{\{g^*(x^*) - f^*(x^*)\}}_{=-\langle x, x^* \rangle}$ and assume that there exists $m_0 \in M$ such that $\|x - m_0\| = \inf_{m\in M} \|x - m\|$. As $x_0^* \in M^\perp$, $\langle m_0, x_0^* \rangle = 0$. Furthermore, since x_0^* maximizes $\langle x, x^* \rangle$, we must have $\|x_0^*\| = 1$.

The second half of the Fenchel duality theorem tells us that

$$f^*(x_0^*) = \max_{y\in X} [\langle y, x_0^* \rangle - \|x - y\|] = \langle m_0, x_0^* \rangle - \|x - m_0\| = -\|x - m_0\|.$$

We have already defined $f^*(\cdot)$ in (6) and know that $f^*(x_0^*) = \langle x, x_0^* \rangle$. Equating these two representations yields

$$\langle x - m, x_0^* \rangle = \langle x, x_0^* \rangle = f^*(x_0^*) = -\|x - m_0\| = -\|x - m_0\| \cdot \|x_0^*\|.$$

This is the alignment requirement except with a sign error. We resolve the sign error by noting that in the Fenchel duality we are using the middle term in (7) while in the minimum norm duality formulation we use the third term in (7), and the difference between the two is a sign.

□

Problem 7.19 (Luenberger)

> Suppose one starts with a fixed quantity x_0 of some resource, which will be allocated at n distinct times as quantities x_1, x_2, and so on. The remaining resource grows by a factor a between allocations. So by necessity, $x_1 \le x_0$, while the second allocation $x_2 \le a(x_0 - x_1)$. Show that a sequence of allocations $\{x_i\}_{i=1}^n$, $x_i \ge 0$ is feasible iff
>
> $$\sum_{i=1}^{n} a^{n-i} x_i \le a^{n-1} x_0. \tag{9}$$
>
> Hence, generalize the result of example 1, section 7.12, to multistage problems.

Solution to 7.19

First we prove that feasibility implies (8). We may extend the chain of restrictions on x_i:

$$x_1 \le x_0; \quad x_2 \le a(x_0 - x_1), \quad x_3 \le a[a(x_0 - x_1) - x_2], \quad \cdots, \quad x_{n-1} \le a^{n-2} x_0 - \sum_{i=1}^{n-2} a^{n-i} x_i,$$

$$\text{and} \quad x_n \le a\left[\left(a^{n-2} x_0 - \sum_{i=1}^{n-2} a^{n-i} x_i\right) - a_{n-1}\right] = a^{n-1} x_0 - \sum_{i=1}^{n-1} a^{n-i} x_i.$$

This last line completes the induction proof, which upon reordering becomes (8).

Second, we show that (8) implies feasibility, i.e. (8) holds for $n = 1, 2, \ldots$ etc. Equation (8) yields

$$a^{n-1}x_0 \geq \sum_{i=1}^{n} a^{n-i}x_i \geq \sum_{i=1}^{k} a^{n-i}x_i,$$

which when divided by a^{n-k} yields our desired result: $a^{k-1}x_0 \geq \sum_{i=1}^{k} a^{k-i}x_i$.

Now we consider a multi-stage allocation problem with investment growth of uncommitted resources. This is simply extending the results of the allocation problem described in example 1, pages 202-203 of Luenberger. We want to maximize $g(x) = \sum_{i=1}^{n} g_i(x_i)$ s.t. $\sum_{i=1}^{n} a^{n-i}x_i = a^{n-1}x_0$ and $x_i \geq 0$. Here the $g_i(\cdot)$'s are increasing concave functions. Let $f(x) = 0$, $D = \{x \mid x_i \geq 0\}$, and

$$C = \left\{ x \;\middle|\; \sum_{i=1}^{n} a^{n-i}x_i = a^{n-1}x_0 \right\} = \left\{ x \;\middle|\; \sum_{i=1}^{n} a^{1-i}x_i = x_0 \right\}.$$

Define $\vec{a} = \left[1, a^{-1}, a^{-2}, \ldots, a^{-(n-1)}\right]^T$. Now we have $C^* = \{y \mid y = \lambda\vec{a}\}$ and $f^*(\lambda\vec{a}) = \lambda x_0$.

$$g_i^*(y_i) = g_i^*\left(\lambda a^{1-i}\right) = \inf_{x_i \geq 0}\left[x_i \lambda a^{1-i} - g_i(x_i)\right]. \tag{10}$$

Further note that $g^*(y) = \sum_{i=1}^{n} g_i^*(y_i)$. By the Fenchel duality theorem, we want to solve

$$\max_{\lambda}\left[g^*(\lambda\vec{a}) - f^*(\lambda\vec{a})\right] = \min_{\lambda}\left[\lambda x_0 - \sum_{i=1}^{n} g_i^*\left(\lambda a^{1-i}\right)\right]. \tag{11}$$

By taking the λ which extremizes (10) and plugging it into (9), we may solve for the x_i which minimize (9). □

Problem 7.20 (Luenberger)

Solution to 7.20

□

Problem 7.21 (Luenberger)

Consider a control system in $\mathbb{R}^n$, governed by the linear equation

$$\dot{x} = Ax + bu,$$

where $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, and $u \in L_2[0, T]$. Using conjugate function theory and duality, find the control u minimizing

$$J = \frac{1}{2}\|x(T)\|^2 + \frac{1}{2}\int_0^T u^2(t)\,\mathrm{d}t.$$

Solution to 7.21

The solution of the matrix differential equation $\dot{x} = Ax + bu$ is

$$x(t) = \mathrm{e}^{At}x_0 + \int_0^t \mathrm{e}^{A(t-\tau)}bu(\tau)\,\mathrm{d}t.$$

Here the matrix exponential e^{At} is defined analogously to a scalar exponential (via infinite summed series): for $c \in \mathbb{R}$, $e^c = 1 + c + c^2/2 + c^3/3 + \cdots = \sum_{k=0}^{\infty} c^k/k$. Similarly, we define the matrix exponential $e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3}A^3 + \cdots$. It is clear that $\frac{d}{dt}e^{At} = Ae^{At}$.

The matrix exponential is the fundamental matrix of solutions in this case, but you may use the notation $\Phi(t)$ in your solutions if you prefer. I choose to be exact in this solution key.

Our strategy is to fix $x_T = x(T)$ and solve the fixed endpoint problem. Then we will minimize with respect to x_T.

We appeal to the Fenchel Duality theorem.[1] Let $u \in L_2[0,1] = C = X$ and $f(u) = \frac{1}{2}\int_0^T u^2(t)\,dt$. Since $X = L_2[0,1]$ is a Hilbert space, we have $X^* = L_2[0,1]$. Now we show that $C^* = L_2[0,1]$:

$$C^* = \left\{ u^* \in L_2[0,1] \;\middle|\; \sup_{u \in L_2} \left[\langle u, u^* \rangle - \frac{1}{2}\int_0^T u^2(t)\,dt \right] < \infty \right\}.$$

As we're in a Hilbert space, $\langle u, u^* \rangle = \langle u \mid u^* \rangle = \int_0^T u(t)u^*(t)\,dt$. By the Cauchy-Schwartz-Buniakovsky inequality, we have $\langle u \mid u^* \rangle \le \|u\| \cdot \|u^*\|$ and we also know that $\int_0^T u^2(t)\,dt = \|u\|^2$, so $\forall u, u^* \in L_2[0,1]$ we have

$$\left[\langle u, u^* \rangle - \frac{1}{2}\int_0^T u^2(t)\,dt \right] \le \|u\| \cdot \|u^*\| - \frac{1}{2}\|u\|^2 < \infty.$$

Choose $g(u) = 0$ and $D = \left\{ u \;\middle|\; \int_0^T e^{A(T-t)} b u(t)\,dt = x_T - e^{AT}x_0 \right\}$. Define the constant $c \equiv x_T - e^{AT}x_0 \in \mathbb{R}^n$ and the linear operator $K : L_2[0,1] \to \mathbb{R}^n$, $Ku = \int_0^T e^{A(T-t)} b u(t)\,dt$; then our notation is simplified as $D = \{u \mid Ku = c\}$. So c and D are functions of x_T. The adjoint of K, K^*, is such that for any $u \in L_2[0,1]$, $a \in \mathbb{R}^n$ we have $\langle a \mid Ku \rangle = \langle K^*a \mid u \rangle$. Expanding the inner products,

$$\langle a \mid Ku \rangle = \underbrace{\left\langle a \;\middle|\; \int_0^T e^{A(T-t)} b u(t)\,dt \right\rangle}_{\text{inner product of vectors}} = \int_0^T a^T e^{A(T-t)} b u(t)\,dt$$

$$= \underbrace{\left\langle b^T e^{A^T(T-t)} a \;\middle|\; u \right\rangle}_{\text{inner product of } L_2 \text{ operators}} = \langle K^*a \mid u \rangle.$$

So $K^* : \mathbb{R}^n \to L_2[0,1]$, $K^*a = b^T e^{A^T(T-t)} a$.

A little thought shows that if $\inf_{u \in L_2[0,1]} \int_0^T u(t)u^*(t)\,dt > -\infty$, then $u^*(t) = K^*a$ for some $a \in \mathbb{R}^n$. Hence $D^* = \{u^* \in L_2[0,1] \mid u^* = K^*a,\ a \in \mathbb{R}^n\}$. The functional $g^* : D^* \to \mathbb{R}$ conjugate to $g = 0$ is such that $g^*(K^*a) = a^T c$. Now we have all the ingredients we need in order to apply Fenchel duality.

$$\inf_{x \in C \cap D} \{f(x) - g(x)\} = \inf_{u \in L_2[0,1]} \left\{ \frac{1}{2}\int_0^T u^2(t)\,dt - 0 \right\} = \max_{\substack{u^* = K^*a \\ a \in \mathbb{R}^n}} g^*(K^*a) - f^*(K^*a)$$

$$= \max_{a \in \mathbb{R}^n} \left\{ a^T c - \frac{1}{2}\int_0^T a^T e^{A(T-t)} b b^T e^{A^T(T-t)} a\,dt \right\}.$$

We will minimize $h(a)$ with respect to a, where

$$h(a) = a^T c - \frac{1}{2}\int_0^T a^T e^{A(T-t)} b b^T e^{A^T(T-t)} a\,dt = a^T c - \frac{1}{2}a^T W a$$

[1] Here we follow the lead of example 3 of page 205-206 of Luenberger

and $W = \int_0^T \mathrm{e}^{A(T-t)} bb^T \mathrm{e}^{A^T(T-t)}\,\mathrm{d}t$. The matrix W has meaning in control theory; it is called the controllability Gramian. W is also symmetric, so $W^T = W$. Taking the derivative with respect to the vector a and setting it to $[0,0,\ldots,0]^T$, we have $\frac{\mathrm{d}}{\mathrm{d}a} h(a) = c - Wa = 0$. So choose any vector a such that $Wa = c$. Then $u = K^* a = b^T \mathrm{e}^{A^T(T-t)} a$.

Let's assume that W is invertible. Then $a = W^{-1}c$, and $u(t) = b^T \mathrm{e}^{A^T(T-t)} W^{-1} c$. Since c depends on x_T, so does $u(t)$. Now we need to minimize $J(x_T) = \frac{1}{2}\|x_T\|^2 + \frac{1}{2}\int_0^T u^2(t)\,\mathrm{d}t$ with respect to x_T. Note that

$$\int_0^T u^2(t)\,\mathrm{d}t = \int_0^T u^T(t) u(t)\,\mathrm{d}t = c^T W^{-1} \underbrace{\int_0^T \mathrm{e}^{A(T-t)} bb^T \mathrm{e}^{A^T(T-t)}\,\mathrm{d}t}_{=W} W^{-1} c = c^T W^{-1} c.$$

Taking the vector derivative and setting it to zero yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x_T} J(x_T) &= \frac{\mathrm{d}}{\mathrm{d}x_T}\left\{\frac{1}{2}\|x_T\|^2 + \frac{1}{2}\left(x_T - \mathrm{e}^{AT}x_0\right)^T W^{-1}\left(x_T - \mathrm{e}^{AT}x_0\right)\right\} \\ &= x_T + W^{-1}\left(x_T - \mathrm{e}^{AT}x_0\right) = \left(I + W^{-1}\right)x_T - W^{-1}\mathrm{e}^{AT}x_0 = 0.\end{aligned}$$

So we get $x_T = \left(I + W^{-1}\right)^{-1} W^{-1} \mathrm{e}^{AT} x_0$. Note that $\left(I + W^{-1}\right)^{-1} W^{-1} = (I + W)^{-1}$, and $x_T = (I + W)^{-1}\mathrm{e}^{AT}x_0$. Since $x_T = -W^{-1}c$, and $u = b^T \mathrm{e}^{A^T(T-t)} W^{-1} c$, we have at last

$$u(t) = -b^T \mathrm{e}^{A^T(T-t)} (I + W)^{-1} \mathrm{e}^{AT} x_0.$$

□

Problem 8.8 (Luenberger)

Solution to 8.8

□

Problem 8.9 (Luenberger)

Show by Lagrange duality that

$$\begin{array}{c} \min\ b^T x \\ Ax \geq c \\ x \geq 0 \end{array} \qquad \text{and} \qquad \begin{array}{c} \max\ c^T \lambda \\ A^T\lambda \leq b \\ \lambda \geq 0 \end{array}$$

are dual problems, where $x, b \in \mathbb{R}^n$, $c, \lambda \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$.

Solution to 8.9

Following the notation of section 8.6 of Luenberger, $f(x) = b^T x$, $G(x) = c - Ax$, and $\Omega = \{x \mid x \geq 0\}$. Define the primal functional $\omega(z) = \inf\{b^T x \mid c - Ax \leq z,\ x \geq 0\}$ and dual functional $\phi(z^*) = \inf_{x \geq 0}\left\{b^T x + z^*(c - Ax)\right\}$. Now the Lagrange duality theorem gives us that

$$\inf_{\substack{c - Ax \leq 0 \\ x \geq 0}} b^T x = \max_{z^* \geq 0} \phi(z^*) = \max_{z^* \geq 0} \inf_{x \geq 0}\left\{b^T x + z^*(c - Ax)\right\} = \max_{z^* \geq 0} \inf_{x \geq 0}\left\{(b^T - z^* A)x + z^* c\right\}.$$

We need that $b^T - z^* A \geq 0$ in order that $\inf_{x \geq 0} \left\{ (b^T - z^* A)x + z^* c \right\} > -\infty$. In that case, we have $\inf_{x \geq 0} \left\{ (b^T - z^* A)x + z^* c \right\} = z^* c$. Then we choose $z^* \geq 0$ which maximizes $z^* c$. Substituting λ^T for z^* gives our desired result. □

Problem 8.10 (Luenberger)

Solution to 8.10

□

Problem 8.14 (Luenberger)

Solution to 8.14

□