Solutions to Homework #7, Math 116

Keziah Cook and Michael McElroy

December 15, 2003

Problem 5.6 (Luenberger)

Consider the scalar system $\ddot{x}(t) = u(t)$, $x(0) = 0$ and $\dot{x}(0) = 1$. We want to choose a control u which minimizes the time to move the system to $x(T) = \dot{x}(T) = 0$. The physical example is a rocket (or car) of unit mass with thrust (acceleration) u. We add the restriction $|u(t)| \leq 1$.

(a) Produce an argument that converts this problem to a minimum norm problem on a time interval $[0, T]$. That is, create a norm problem in terms of u, and then relate the solution's u to the final time T.

(b) Solve the problem.

Solution to 5.6

Obviously, we can integrate to get constraints based on our final conditions:

$$\begin{aligned}
\dot{x}(T) &= 1 + \int_0^T u(\tau)\,\mathrm{d}\tau = 0; \\
x(T) &= \int_0^T 1 + \int_0^\sigma u(\tau)\,\mathrm{d}\tau\,\mathrm{d}\sigma = T + \underbrace{\left\{\int_0^\tau u(\sigma)\,\mathrm{d}\sigma \cdot \sigma\right\}\Big|_0^T}_{=-T} - \int_0^T \tau u(\tau)\,\mathrm{d}\tau \\
&= -\int_0^T \tau u(\tau)\,\mathrm{d}\tau = 0.
\end{aligned}$$

Above we integrated by parts and inserted our knowledge of $\dot{x}(T)$ to simplify the equation for $x(T)$. So our three constraints come from our velocity and position endpoints and our magnitude limitation:

$$\int_0^T u(\tau)\,\mathrm{d}\tau = -1; \tag{1}$$

$$\int_0^T \tau u(\tau)\,\mathrm{d}\tau = 0; \tag{2}$$

$$|u(t)| \leq 1. \tag{3}$$

Change the problem

We cannot directly solve for the minimal time to move the system to rest at the origin such that $|u(t)| \leq 1$. We can, however, ask if for any given T there exists a u that fulfills (1), (2), and (3). If so, then $T_{min} \leq T$, while if not, then $T_{min} > T$.

Let's ignore the constraint on the control for a moment. Then we could move the system to rest at the origin at an arbitrarily small time. For instance, if we had infinite thrust, we could set $u(t) = -\delta(t)$, with $\delta(t)$ being the Dirac delta function, and we'd have stopped the system at time 0! Intuitively, if we wanted to get to $x(t) = \dot{x}(T) = 0$ for small T, we would have to use large $|u(t)|$, but if we want to bring the system to rest in after a long time, say, $T = 100$, then we could slow down the system leisurely with a thrust u of small magnitude.

So, without the constraint $|u| \leq 1$, we can find a thrust schedule $u(t)$ such that conditions (1) and (2) are met (i.e. $\dot{x}(T) = x(T) = 0$) for any positive T.

We change our problem. We ask, for a given end time T, what are the possible thrust schedules $u_i(t)$ that fulfill conditions (1) and (2). Call the set of these $M = \left\{ u_i \;\middle|\; \int_0^T u(\tau)\,\mathrm{d}\tau = -1, \int_0^T \tau u(\tau)\,\mathrm{d}\tau = 0 \right\}$. We ask if there is a thrust $u_i(t) \in M$ such that $|u_i(t)| \leq 1$ for all $t \in [0, T]$. If there is, then $T_{min} \leq T$. We need only consider those $u_i \in M$ such that $\|u_i\|_\infty = \min_{u \in M} \|u\|_\infty$. So, we hit upon an idea: for any given T, we'll solve for the minimal value of $\|u\|_\infty = \sup\limits_{t \in [0,T]} |u(t)|$ in M. Call this minimal norm $n(T)$. Note that $n(T) > 0$ and is a non-increasing function, i.e. $T_1 < T_2 \implies n(T_1) \geq n(T_2)$. We simply find where $n(T)$ first equals 1; that is our T_{min}.

To solve our new problem, we will apply corollary 1 from page 123 of Luenberger:

Corollary. *Let $y_i \in X$, and suppose there exists a set $S \subset X^*$ s.t. all $x^* \in S$ fulfill the system of linear equalities $\langle y_i, x^* \rangle = c_i$. Then we have*

$$\min_{x^* \in S} \|x^*\|_{X^*} = \max_{\left\|\sum_i a_i y_i\right\|_X \leq 1} \sum_i a_i c_i, \quad a_i \in \mathbb{R}.$$

Furthermore, the optimal x^* is aligned with the optimal $\sum_i a_i y_i$.

We need to choose a space for u to live in which is the dual of another space. We are already considering minimizing the sup-norm of u. This is handy since then we consider $u(t) \in L_\infty[0, T]$, and $L_\infty = (L_1)^*$. We will solve our problem the same as Example 2 from section 5.9 of Luenberger.

Given $T > 0$, we have $X = L_1[0, T]$, $X^* = L_\infty[0, T]$, and conditions (1) and (2) yield $y_1(t) = 1$, $c_1 = -1$, $y_2(t) = t$, $c_2 = 0$. Our optimization problem is

$$\min_{\substack{\langle 1, u \rangle = -1 \\ \langle t, u \rangle = 0}} \|u\|_\infty = \max_{\|a_1 + a_2 t\|_1 \leq 1} -a_1.$$

We work on the maximization problem. Note that we have gone from an infinite-dimensional problem, choosing the curve u, so the two-dimensional problem of finding $a_1, a_2 \in \mathbb{R}$. It's a miracle! We do not even have to solve the maximization problem, because the alignment insight is enough to allow us to solve the minimization problem! $a_1 + a_2 t$ is a line, so the optimal u must be aligned with a line. We need to state what it means to be aligned.

Alignment theorems

Definition. A vector $x^* \in X^*$ is aligned with a vector $x \in X$ iff $\langle x, x^* \rangle = \|x^*\|_{X^*} \|x\|_X$.

That is, if we have a normed linear space M, then $x \in M$ and $y \in M^*$ are aligned if $y(x) = \|x\|_M \|y\|_{M^*}$.

Theorem. *Let $X = L_p[a, b]$, $1 < p < \infty$, and $X^* = L_q[a, b]$, $\frac{1}{p} + \frac{1}{q} = 1$. Then $x \in L_p$, $y \in L_q$ are aligned iff $x(t) = K \cdot [sgn\ y(t)] \cdot |y(t)|^{q/p}$.*

Proof. We have $\langle x, y\rangle = \langle x|y\rangle$, so the functional evaluation is also the inner product. Inserting this into the Hölder inequality yields

$$\langle x, y\rangle \leq |\langle x, y\rangle| \leq \|x\|_p \|y\|_q$$

with equality (of the middle and right terms) iff $x(t)$ and $y(t)$ are proportional. We add the sign requirement to get equality of the left and middle terms. □

Corollary (Alignment between L_1 and L_∞). *If $x \in L_\infty[a, b]$ and $y \in L_1[a, b]$ are aligned, then we have $x(t) = K \cdot sgn\ y(t)$.*

Theorem. *Let $x \in X = C[a, b]$ and let $\Gamma = \{t \in [a, b] \,|\, |x(t)| = \|x\| = \max |x|\}$. A bounded linear functional $x^*(x) = \int_a^b x(t)\mathrm{d}v(t)$ is aligned with x iff v varies only on Γ and v is non-increasing at t if $x(t) < 0$ and non-decreasing at t if $x(t) > 0$.*

Solving our new problem

Applying our alignment corollary, we have $u(t) = \pm K$, for some $K \in \mathbb{R}$. Furthermore, since a line may change signs at most once, so does u. It is obvious that we will choose $u < 0$ at first, to slow down our rocket and move back to the origin, but then we have to apply brakes ($u > 0$) in order to stop to rest at $x = 0$. So we have only two parameters to solve for, the switch time s and the magnitude K of the thrust u, and we choose s and K so as to obey the constraints (1) and (2). So $u = \begin{Bmatrix} -K & t \in [0, s) \\ K & t \in [s, T] \end{Bmatrix}$. Thus we are writing $K = \|u\|_\infty$. Our velocity constraint (1) yields

$$-1 = \int_0^T u\mathrm{d}t = \int_0^s -K\mathrm{d}t + \int_s^T K\mathrm{d}t = K(T - 2s) \implies K = \frac{1}{2s - T}.$$

Our position constraint (2) yields

$$0 = \int_0^T tu(t)\,\mathrm{d}t = -K\int_0^s t\,\mathrm{d}t + K\int_s^T t\,\mathrm{d}t = \frac{K}{2}(T^2 - 2s^2) \implies s = \frac{\sqrt{2}}{2}T.$$

So $KT = \sqrt{2} + 1$ and $K = \|u\|_\infty$. We see that we may find a u fulfilling $\|u\|_\infty \leq 1$, i.e. constraint (3), for all $T \geq \sqrt{2} + 1$, so our minimum time is $\boxed{T_{min} = \sqrt{2} + 1}$. The curves for $x(t)$ and $\dot{x}(t)$ are in figure 1.

Generalizing the problem

Figure 2 is far more interesting. It show the effect in phase-space of applying thrusts of ± 1. If $u = +1$, one moves along a blue (right-facing) parabola, while if $u = -1$, then one moves along a red parabola. These parabolas have the form $x = \pm\frac{1}{2}\dot{x}^2 + c$. Time optimal strategies are obtained by following these parabolas, with at most one switch. This solves the more general problem of finding the time optimal thrust pattern to get from $(x, \dot{x})(0) = (a, b)$ to $(x, \dot{x})(T) = (c, d)$.

The equation $KT = \sqrt{2} + 1$ is better written as $(\min \|u\|_\infty) \cdot T_{min} = \sqrt{2} + 1$. This shows that the minimum time to bring the system to rest at the origin depends inversely on the bound on the sup-norm of our thrust. So if would allow thrusts as big as ± 2, then we could halve the time necessary to return the system to rest at the origin.

We may also solve the dual problem and show that it gives the same maximum as the minimum. We want a_1, a_2 which maximize $-a_1$ while having $\|a_1 + a_2 t\|_1 \leq 1$. Plotting the line $y(t) = a_1 + a_2 t$ along the time axis t, we have that a_1 is the y-intercept. Figure 3 shows the graph of $a_1 + a_2 t$; we want to decrease the y-intercept while maintaining the area of the shaded region at 1. We realize that we want a

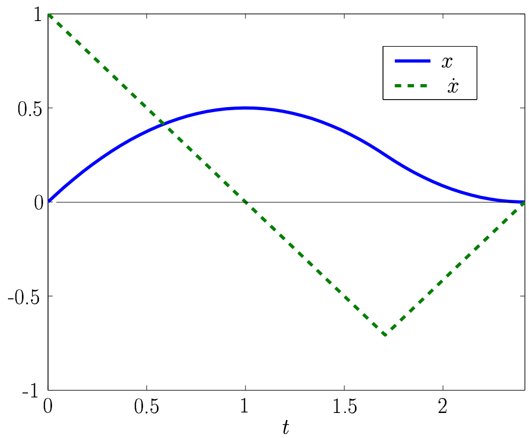

Figure 1: Curves for the state variables x and $\dot{x}$. Note that the thrust function is $u = -1$ until $t = 1 + \frac{\sqrt{2}}{2}$, whereafter $u = 1$.

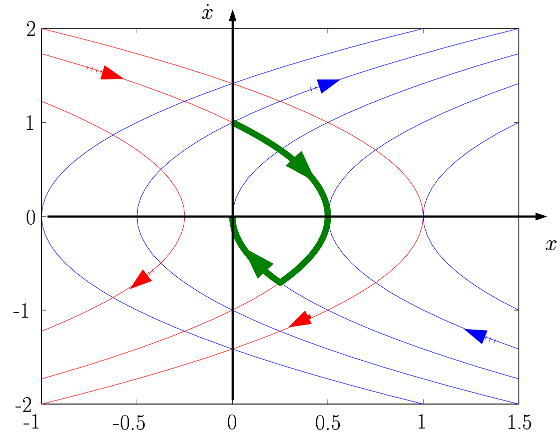

Figure 2: Phase plane portrait for the family of optimal trajectories. The bold (green) curve is the plot of figure 1 in these new coordinates.

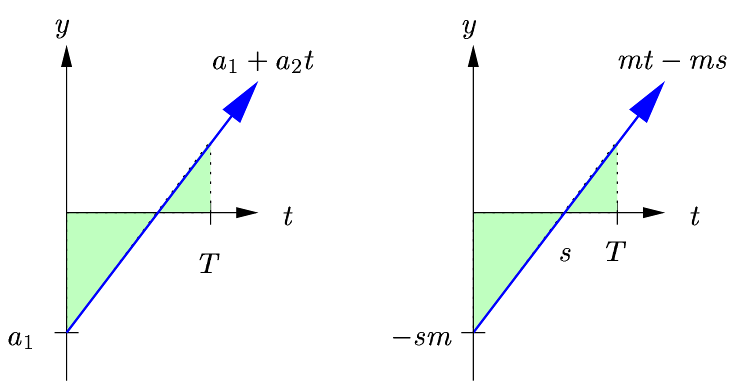

Figure 3: For the dual maximization problem, here are graphed the vector $a_1 + a_2 t \in L_1[0, T]$ and the constraint "shaded area" $= \|a_1 + a_2 t\|_1 \leq 1$. The figure on the right is the same except written with $s = a_1/a_2$ and $m = a_2$

negative y-intercept (so that $-a_1 > 0$) and positive slope. We choose to change coordinates to simplify the analysis: let $y(t)$ have x-intercept s (the switch time) and slope m. Then the y-intercept is $-sm$ while the area is $\frac{m}{2}\left[s^2 + (T-s)^2\right]$. So we have

$$\max_{\|a_1+a_2 t\|_1 \leq 1} -a_1 = \max_{\|m(t-s)\|_1 \leq 1} sm,$$

and we set $m = \frac{2}{T^2-2sT+2s^2}$. Now we want to maximize $\frac{2s}{T^2-2sT+2s^2}$ with respect to s, T. Taking the partial wrt s and setting to zero yields $0 = \frac{2(T^2-2sT+2s^2)-2s(-2T+4s)}{(T^2-2sT+2s^2)^2}$, or $s = \frac{\sqrt{2}}{2}T$. Hence $sm = \frac{\sqrt{2}+1}{T}$. If this is our maximum to the dual problem, it is the minimum to the primal problem, so $\min \|u\|_\infty = \frac{\sqrt{2}+1}{T}$, which is what we worked out earlier. □

Problem 5.10 (Luenberger)

Let X be a normed space and M a subspace of it. Show that (within an isometric isomorphism) $M^\star = X^\star/M^\perp$ and $M^\perp = (X/M)^\star$.

Solution to 5.10

Claim. $M^\star = X^\star/M^\perp$

Proof. Define the mapping $T : M^\star \mapsto X^\star/M^\perp$ as follows. Let $f \in M^\star$ be a bounded linear functional on M. By the Hahn-Banach Theorem, we can extend f to a linear functional F on X. Define $T(f) = [F] \in X^\star/M^\perp$. We want to show that T is an isometric isomorphism.

1. T is well-defined. Let F_1, F_2 be two possible extensions of f. Then

$$\begin{aligned} F_1(m) &= F_2(m) \quad \forall m \in M \\ (F_1 - F_2)(m) &= 0 \quad \forall m \in M \end{aligned}$$

   Thus $(F_1 - F_2) \in M^\perp$ and $[F_1] = [F_2]$. So $T(f) = [F]$ doesn't depend on the choice of extension F.

2. T is Linear

$$\begin{aligned} T(\alpha f + \beta g) &= [\alpha F + \beta G] \\ &= \alpha[F] + \beta[G] \\ &= \alpha T(f) + \beta T(g) \end{aligned}$$

3. T is injective. Suppose $f_1, f_2 \in M^\star$ are such that $f_1(m) \neq f_2(m)$ for some $m \in M$. Let $T(f_1) = [F_1], T(f_2) = [F_2]$. Note that $(F_1 - F_2)(m) = f_1(m) - f_2(m) \neq 0$ for some $m \in M$. Thus $(F_1 - F_2) \notin M^\perp$ and $[F_1] \neq [F_2]$.

4. T is surjective. Let $F_1, F_2 \in [F]$. Then $F_1(m) = F_2(m) \forall m \in M$. So $F_1|_M = F_2|_M = f$. (The point here is that restriction is well-defined). Clearly, $T(f) = [F]$ so T is surjective.

5. T is an isometry. Let $T(f) = [F]$ and note that $\|f\|_M = \|F\|$, when F is chosen to be the minimum norm extension of f. (This exists by the Hahn-Banach theorem). This means that:

$$\sup_{\|x\|=1, x\in M} | f(x) | = \sup_{\|x\|=1} |F(x)|$$

But for any $(F+m)$ where $m \in M^\perp$, we have

$$\begin{aligned}
\|F+m\| &= \sup_{\|x\|=1} |(F+m)x| \\
&\geq \sup_{\|x\|=1 m x\in M} |(F+m)x| \\
&\geq \sup_{\|x\|=1, x\in M} |F(x)| \\
&= \sup_{\|x\|=1, x\in M} |f(x)| \text{ , because } F \text{ agrees with } f \text{ on } M \\
&= \|f\|_M = \|F\|.
\end{aligned}$$

Thus $\inf_{m\in M^\perp} \|F+m\| \geq \|F\|$, but note that this lower bound is reached at $m = \theta \in M^\perp$. Thus $T(f) = \|[F]\| = inf_{m\in M^\perp}\|F+m\| = \|f\|_M$. And T is an isometry.

Thus T is an isometric isomorphism. □

Claim. $M^\perp = (X/M)^\star$

Proof. Given a functional $f \in M^\perp$ defined on X, define the functional $\overline{f}$ on X/M by $\overline{f}([x]) = f(x)$. To see that this is well-defined, suppose $x_1 \neq x_2$ but $[x_1] = [x_2]$, i.e. $(x_1 - x_2) \in M$. Then $\overline{f}([x_1]) - \overline{f}([x_1]) = f(x_1) - f(x_2) = f(x_1 - x_2) = 0$ because $(x_1 - x_2) \in M$ and $f \in M^\perp$. Thus the definition of $\overline{f}$ doesn't depend on the representation of its argument.

Let $T(f) = \overline{f}$ as above. If f_1, f_2 are distinct, they differ for some $x \notin M$. $f_1(x) \neq f_2(x) \implies \overline{f}_1([x]) \neq \overline{f}_2([x])$. Thus T is one-to-one.

Define the inverse map $T^{-1} : (X/M)^\star \mapsto M^\perp$ by $T^{-1}(\overline{f}) = f$ where $f(x) = \overline{f}([x])$ for all $x \in X$. This is well-defined and injective by the above arguments. The range is indeed $M^\perp$ because for $m \in M, f(m) = \overline{f}([m]) = \overline{f}([0]) = 0$, because $\overline{f}$ is linear.

Finally, note that for any element $(x+m) \in [x]$, we have $f(x+m) = f(x) + f(m) = f(x)$. Now,

$$\|T(f)\| = \sup_{\|[x]\|=1} |\overline{f}([x])| = \sup_{\|[x]\|=1} |f(x)|$$

Now there might be an x such that $\|x\| \neq 1$ but $\inf_{m\in M} \|x+m\| = \|[x]\| = 1$. Thus there is a sequence m_i such that $\|x + m_1\| \to 1$. Note that $f(x+m_i) = f(x)$ for all i, so it is sufficient to consider representation elements with $\|x\| = 1$ in the equation above. Thus

$$\|T(f)\| = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\|\leq 1} |f(x)|| = \|f\|$$

Thus we see that T is an isometric isomorphism. □

□

Problem 5.15 (Luenberger)

Let X be a real linear vector space and let $f_1, \ldots, f_n$ be linear functionals on X. Show that for any given $\{\alpha_i\}$, the following statements are equivalent:

1. $\exists x \in X, f_i(x) = \alpha_i$ for $i = 1, \ldots, n$.
2. $\displaystyle\sum_{i=1}^n \lambda_i f_i = 0 \implies \sum_{i=1}^n \lambda_i \alpha_i = 0.$

Solution to 5.15

$1 \implies 2$: Assume we have an $x \in X$ s.t. for all i we have $f_i(x) = \alpha_i$. Then we have

$$\sum_{i=1}^{n} \lambda_i f_i = 0 \implies \sum_{i=1}^{n} \lambda_i \underbrace{f_i(x)}_{=\alpha_i} = 0 \implies \sum_{i=1}^{n} \lambda_i \alpha_i = 0.$$

$2 \implies 1$: First, we define a function $F : X \to \mathbb{R}^n$, where $F(x) = [f_1(x), f_2(x), \ldots, f_n(x)]^T$. First we show that $M = \text{image}(F)$ is a subspace of $\mathbb{R}^n$. Let $y_1, y_2 \in M$, $\alpha, \beta \in \mathbb{R}$ be given. Choose $x_1, x_2 \in X$ s.t. $F(x_1) = y_1$, $F(x_2) = y_2$. Then the linearity of the f_i functionals gives us that $\alpha y_1 + \beta y_2 = \alpha F(x_1) + \beta F(x_2) = F(\alpha x_1 + \beta x_2)$, and so $\alpha y_1 + \beta y_2 \in M$. Furthermore, M is nonempty since $\theta \in X$, so $F(\theta) = 0 \in M$. M is closed, since we are in a finite-dimensional space $\mathbb{R}^n$.

Now we consider using vectors for our statements: we want to show that if $\forall x \in X, \lambda^T F(x) = 0 \implies \lambda^T \alpha$ then we have some $x \in X$ with $F(x) = \alpha \in \mathbb{R}^n$.

Idea 1 Assume statement 2. Consider those $\lambda \in \mathbb{R}^n$ such that $\forall x \in X, \lambda^T F(x) = 0$. This means that $\lambda \perp M$, or $\lambda \in M^\perp$. If $\lambda^T \alpha = 0$, then $\lambda \perp \alpha$. So we have α is orthogonal to all vectors in $M^\perp$. Hence $\alpha \in {}^\perp[M^\perp]$, ${}^\perp[M^\perp] = M$ by theorem 1 of page 118 in Luenberger. So there exists some $x \in X$, $F(x) = \alpha$.

Idea 2 Use the geometric Hahn-Banach theorem. Assume statement 2. If $\alpha \in M$, then we are done. Assume $\alpha \notin M$; we'll show a contradiction. If M has no interior, then either $M = \{0\}$ or $M = \{cy\}$ for some fixed $y \in \mathbb{R}^n$. If $M = \{0\}$, then $f_i = 0$. Then all $\lambda \in \mathbb{R}^n$ are such that $\sum_i \lambda_i f_i = 0$, and hence we must have $\alpha_i = 0$. So $\alpha \in M$. If M is one-dimensional, then an argument like Idea 1 says that $\alpha \in M$.

If, however, M has an interior, then we may choose a hyperplane H separating M and α: $H = \{x \,|\, p^T x = c\}$, and $\forall m \in M$, we have $p^T m < c < p^T \alpha$. Here p is a normal vector to H, $p \in (\mathbb{R}^n)^* = \mathbb{R}^n$.

Since M is a subspace, $0 \in M$, and hence $0 = p^T 0 < c < p^T \alpha$. Furthermore, H and M do not intersect, so they must be parallel. Hence $p^T m_1 = p^T m_2$ for all $m_1, m_2 \in M$, i.e. $p^T m = 0$ for all $m \in M$. So we may consider $\lambda = p$. Then for all $x \in X$, $\lambda F(x) = p^T F(x) = 0$, while $\lambda^T \alpha > c > 0$. So we have a contradiction.

□

Problem 5.16 (Luenberger)

Let $g_1, \ldots, g_n$ be linearly independent functionals a vector space X. Let f be another linear functional on X such that for every $x \in X$ satisfying $g_i(x) = 0$ for each i, we have $f(x) = 0$. Show that there are constants $\lambda_1, \ldots, \lambda_n$ such that

$$f = \sum_{i=1}^{n} \lambda_i g_i$$

.

Solution to 5.16

We will first prove that the statement holds if X is finite dimensional. We will then use this lemma to prove the statement in generality.

Lemma. *The problem statement holds if X is finite-dimensional.*

Proof. It suffices to show the result in the case where $X = R^m$ where m is a nonnegative integer (as all m-dimensional real vector spaces are isomorphic to R^m). Since R^m is isomorphic to $(R^m)^\star$, there exists a vector $x_i^\star \in R^m$ such that for each $x \in X$, $g_i(x) = \langle x_i^\star \mid x \rangle$ Similarly, there exists a vector $y^\star \in R^m$ such that for each $x \in X$, $f(x) = \langle y^\star \mid x \rangle$.

Let $S = [\{x_1^\star, \dots, x_n^\star\}]$. Let K_i be the kernel of g_i, i.e. $K_i = \{x \in R^m \mid g_i(x) = 0\}$. Let K be the intersection of the K_i's. We note that K is a subspace of X, as the g_i are linear functionals. We want to show that $S = K^\perp$, as this is equivalent to the problem statement. Since R^m is finite dimensional, all subspaces are closed. Thus $S^\perp = K$ implies $S = K^\perp$.

- $K \subset S^\perp$

  Let $x \in K$. Fix an arbitrary $s \in S$. As S is the span of $\{x_1^\star, \dots, x_n^\star\}$, we can write $s = \alpha_1 x_1^\star + \cdots + \alpha_n x_n^\star$.

$$\begin{aligned} \langle s \mid x \rangle &= \langle \alpha_1 x_1^\star + \cdots + \alpha_n x_n^\star \mid x \rangle \\ &= \alpha_1 \langle x_1^\star \mid x \rangle + \cdots + \alpha_n \langle x_n^\star \mid x \rangle \\ &= \alpha_1 g_1(x) + \cdots + \alpha_n g_n(x) \\ &= 0 \end{aligned}$$

  because $x \in K$. As our choice of s was arbitrary, we see $x \in S^\perp$.

- $S^\perp \subset K$

  Suppose $x \in S^\perp$. This means $\langle x_i^\star \mid x \rangle = 0$ for all i. By our choice of $x_i^\star$, $g_i(x) = \langle x_i^\star \mid x \rangle = 0$. Thus $x \in K_i$ for all i, and we see $x \in K$.

Thus we see that $S^\perp = K$. Since S, K are closed subspaces, we know that $S = S^{\perp\perp} = K^\perp$.

Functions in $K^\perp$ are exactly those functions $f(x)$ such that $f(x) = 0$ whenever $g_i(x) = 0, \forall i$. We just showed that any such f is in the span of $\{x_i^\star, \dots, x_n^\star\}$ and thus can be written as $f = \sum_{i=1}^n \lambda_i g_i$. □

Now we will prove the infinite dimensional case. Let X be a real vector space. Let $g_1, \dots, g_n$ be linearly independent linear functionals on X and let f be another linear functional on X such that for every $x \in X$ satisfying $g_i(x) = 0, i = 1, \dots, n$, we have $f(x) = 0$. As before, define the subspace K_i of X to be the kernel of g_i and define $K = K_1 \cap K_2 \cap \cdots \cap K_n$; note that f is zero on K.

Observe that for each i, K_i has codimension at most 1 (because the range of g_i is the real numbers, which is a one-dimensional vector space); this implies that K has codimension at most n. Therefore, X/K is a finite-dimensional vector space.

For $i = 1, \dots, n$, define $g_i' : X/K \mapsto R$ as follows: for each $\overline{v} \in X/K$, if w is an element of $v + K$ then $\hat{(g)}_i(\overline{v}) = g_i(w)$. This map is well-defined because if w_1 and w_2 are elements of the same equivalence class in X/K then $g(w_1) - g(w_2) = g(w_1 - w_2) = 0$ (because $w_1 - w_2 \in K \subset K_i$); further, it is linear because if $\overline{v}_1$ and $\overline{v}_2$ are elements of X/K, a_1 and a_2 are real scalars, and w_1 and w_2 are representatives of the class $\overline{v}_1$ and $\overline{v}_2$ respectively, then $a_1 w_1 + a_2 w_2$ is a representative of the class $a_1 \overline{v}_1 + a_2 \overline{v}_2$, so

$$\begin{aligned} \hat{g}_i(a_1 v_1 + a_2 v_2) &= g_i(a_1 w_1 + a_2 w_2) \\ &= a_1 g_i(w_1) + a_2 g_i(w_2) = a_1 \hat{g}_i(\overline{v}_1) + a_2 \hat{g}_i(\overline{v}_2) \end{aligned}$$

Similarly, define $\hat{f} : X/K \mapsto R$ as follows: for each $\overline{v} \in X/K$, if w is an element of $v + K$ then $\hat{f}(\overline{v}) = f(w)$. The fact that $\hat{f}$ is well-defined and linear follows from the above argument.

Now suppose $\hat{g}_i(\overline{v}) = 0$ for $i = 1, \ldots, n$. Letting v be a representative of $\overline{v}$, it follows that $g_i(v) = 0$ for all such i; therefore, $f(v) = 0$ because f is 0 whenever all the g_i's are 0, so $\hat{f}(v) = 0$ by the definition of f. Therefore, for every $\overline{v} \in X/K$ satisfying $\hat{g}_i(v) = 0, i = 1, 2, \ldots, n$, we have $\hat{f}(\overline{v}) = 0$. It follows by the lemma that there exists constants $\lambda_1, \ldots, \lambda_n$ such that

$$\hat{f} = \sum_{i=1}^{n} \lambda_i \hat{g}_i$$

Now suppose $v \in X$ and let $\overline{v}$ be the coset in X/K to which v belongs. Therefore,

$$\begin{aligned} f(v) &= \hat{f}(\overline{v}) \\ &= \sum_{i=1}^{n} \lambda_i \hat{g}_i(\overline{v}) \\ &= \sum_{i=1}^{n} \lambda_i g_i(x) \end{aligned}$$

This holds for each $x \in X$. It follows that $f = \sum_{i=1}^{n} \lambda_i g_i$.

□

Problem 5.23 (Luenberger)

> Let X be a real normed linear space, and let K be a convex set in X, having θ as an interior point. Let h be the support functional of K and define $K^\circ = \{x^\star \in X^\star : h(x^\star) \leq 1\}$. Now for $x \in X$, let $p(x) = \sup_{x^\star \in K^\circ} \langle x, x^\star \rangle$. Show that p is equal to the Minkowsik functional of K.

Solution to 5.23

To show that p as defined to this problem is equal to the Minkowski functional q of K, it suffices to show that for each $x \in X$, $p(x) = q(x)$, that is $p(x) \leq q(x)$ and $p(x) \geq q(x)$.

Suppose $x \in X$ and fix $\epsilon > 0$. It follows that there is some $y \in K^0$ such that $\langle x, y \rangle > p(x) - \epsilon$. As $h(y) \leq 1$, it follows that for each $k \in K$, $\langle k, y \rangle \leq h(y) \leq 1$ by the definition of the support functional. Therefore, if r is positive with $r < p(x) - \epsilon$, then

$$\left\langle \frac{x}{r}, y \right\rangle > \frac{(p(x) - \epsilon)}{r} > 1$$

so $\frac{x}{r} \notin K$ for any such r, and this implies that $q(x) \geq p(x) - \epsilon$ for all positive ϵ. Thus $q(x) \geq p(x)$ which implies that $p(x) \leq q(x)$.

Now, if $x = \theta$ then clearly $q(x) = 0 = p(x)$. Therefore, to prove that $p(x) \geq q(x)$ we can assume $x \neq \theta$. In this case, it follows by Corollary 2 of Section 5 that there exists a nonzero $y \in X^*$ such that $|\langle x, y \rangle| = ||x|| \cdot ||y||$. If $h(y) = 0$ then $y \in K$ so

$$p(x) \geq ||x|| \cdot ||y|| > 0$$

However, if $h(y) \neq 0$, then $h(y) > 0$. Setting $y' = \frac{y}{h(y)}$, we have $h(y') = 1$ so that $y' \in K$ and

$$p(x) \geq \langle x, y' \rangle = \left\langle x, \frac{y}{h(y)} \right\rangle = ||x|| \cdot \frac{||y||}{h(y)} > 0$$

This implies that in any event, $p(x) > 0$. Fix $\epsilon > 0$. Letting $z = \frac{x}{p(x)+\epsilon}$ in this case, we seek to show $z \in K$ as that would imply that $q(x) \leq p(x) + \epsilon$. Assume the contrary. By Theorem 3 of Section 5.12, there exists an element $w \in X^*$ such that

$$sup\{\langle x, w \rangle \mid x \in K\} \leq \langle z, w \rangle$$

(as the single point set z is convex). Because $h(\lambda t) = \lambda h(t)$ whenever λ is a positive real and t is an element of X^*, w can be taken so that $h(w) = 1$. By the definition of the support functional, this implies that $1 = h(w) \leq \langle z, w \rangle$ so

$$\langle x, w \rangle = (p(x) + \epsilon) \langle z, w \rangle > p(x) + \frac{\epsilon}{2}$$

However, as $w \in K^0$ because $h(w) = 1$, we know that

$$p(x) = sup\{\langle x, w \rangle \mid w \in K^0\} \geq p(x) + \frac{\epsilon}{2}$$

producing a contradiction. Consequently, $z \in K$ and $q(x) \leq p(x) + \epsilon$ for each $\epsilon > 0$, so that $p(x) \geq q(x)$ for each nonzero x (and thus for all $x \in X$).

Therefore, for all $x \in X$, $p(x) = q(x)$, which implies that $p = q$. □

Problem 6.12 (Luenberger)

Let X, Y be Banach spaces and let $A \in B(X, Y)$ have closed range. Show that

$$\inf_{Ax=b} \|x\|_X = \max_{\|A^*y^*\|_{X^*} \leq 1} \langle b, y^* \rangle.$$

Use this result to reinterpret the solution of the rocket problem of example 3, section 5.9 of Luenberger.

Solution to 6.12

Let $b \in Y$ be given; let $\overline{x}$ be given such that $A\overline{x} = b$. Define $M = \mathcal{N}(A)$; M is a subspace of $\mathcal{N}(A)$. Since A is bounded, A is continuous, and hence M is closed. Then $M^\perp = [M]^\perp = [\mathcal{N}(A)]^\perp = \mathcal{R}(A^*)$ by theorem 2 of page 156 of Luenberger. We have

$$\inf_{Ax=b} \|x\|_X = \inf_{m \in M} \|\overline{x} - m\|_X = \max_{\substack{\|x^*\|_{X^*} \leq 1 \\ x^* \in M^\perp}} \langle x, x^* \rangle.$$

The last equality comes from theorem 1 of page 119 of Luenberger. For any $x^* \in M^\perp = \mathcal{R}(A^*)$, we may choose $y^* \in Y^*$ such that $x^* = A^*y^*$. Thus we have

$$\inf_{Ax=b} \|x\|_X = \max_{\|A^*y^*\|_{X^*} \leq 1} \langle b, y^* \rangle.$$

Finally, we have $\langle x, A^*y^* \rangle = \langle Ax, y^* \rangle = \langle b, y^* \rangle$, which yields our desired result

$$\inf_{Ax=b} \|x\|_X = \max_{\|A^*y^*\|_{X^*} \leq 1} \langle b, y^* \rangle.$$

Let's translate the rocket problem into this notation. $X = NBV[0, T]$, $Y = \mathbb{R}$. For $v \in X$, $Ax \equiv \int_0^T (T - t) dv(t)$; $b = 1 + \frac{T^2}{2}$. $Y^* = \mathbb{R}$, so A^* is simply multiplication by the function $T - t \in C[0, T]$. We know that $(C[0, T])^* = NBV[0, T]$. We need to show that our A here has closed range to apply the result of this problem, and we leave that to the reader. □

Problem 6.14 (Luenberger)

Consider inequalities component-wise. Prove the Minkowski-Farkas lemma, and give a geometric interpretation.

Proposition. *Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$ be given. The following statements are equivalent:*

$$1. \quad Ax \leq 0 \implies b^T x \leq 0.$$
$$2. \quad \exists \lambda \geq 0, \quad b = A^T \lambda.$$

Solution to 6.14

$2 \implies 1$: Let $b = A^T\lambda$, where $\lambda_1, \ldots, \lambda_n \geq 0$. Then we have $0 \geq Ax \implies 0 \geq \lambda^T Ax = b^T x$.

$1 \implies 2$: Assume that $Ax \leq 0 \implies b^T x$, and assume $b \neq A^T\lambda$, $\lambda \geq 0$. Then $b \notin \text{cone}(A)$. If cone(A) is a ray, i.e. rank$(A) = 1$, then let A_1 be the first row of A (all rows are dependent). Consider the half-space $M = \{x, A_1 x \leq 0\}$. This is the set of vectors such that $Ax \leq 0$. If we have $x \in M \implies b^T x = x^T b \leq 0$, then $b = \alpha A_1^T$ for some $\alpha \geq 0$. Note that b could be the zero vector, in which case the statement is trivial.

If rank$(A) > 1$, then cone(A) has an interior. We apply the separating hyperplane theorem, choosing a $p \in \mathbb{R}^n \ \{0\}$, $c \in \mathbb{R}$ s.t. $\forall x \in \text{cone}(A)$ we have $p^T x < c < p^T b$. We need a lemma:

Lemma. *Let K be a convex cone in $\mathbb{R}^n$ and let $p \in \mathbb{R}$. If $p^T x$ is bounded above for all $x \in K$, then $p^T x \leq 0$ for all $x \in K$.*

Proof. We'll use our cone generated by A as our example. We note that the rows of A are in cone(A), so consider $p^T A_i^T < c$. If for any row i we have $p^T A_i^T = d > 0$, then we may find a "scaled up" vector in cone(A), $y = \left(\frac{p^T b}{d} + \epsilon\right) A_i$ such that $p^T y = p^T b + \epsilon d$ (here $\epsilon > 0$). But we have that $p^T x < p^T b$ for all x in the cone, so we must have that $p^T A_i^T \leq 0$. □

By our lemma we have $p^T A_i^T \leq 0 < c < p^T b$. Note that $p^T A_i \leq 0 \implies Ap \leq 0$, so p is am example of an x s.t. $Ax \leq 0$. However, we have $p^T b = b^T p > 0$, so we have a contradiction.

□

Problem 6.15 (Luenberger)

Let M be a closed subspace of a Hilbert space H. Consider the projection operator P, which projects onto M. Show that the projection operator P is linear and bounded with $\|P\| = 1$ if M is at least one-dimensional.

Solution to 6.15

Assume $M \neq \{0\}$ (i.e. M is at least one-dimensional), and that M is closed. We want to show that for all $\alpha, \beta \in \mathbb{R}$, $x, y \in M$, we have $P(\alpha x + \beta y) = \alpha P(x) + \beta P(y)$ (linearity), and that there exists $K > 0$ s.t. $\forall x \in M$, $P(x) \leq K\|x\|$ (boundedness).

Let $x, y, \in M$ be given. Express x, y in terms of their M and $M^\perp$ components: $x = x_M + x_{M^\perp}$ and $y = y_M + y_{M^\perp}$. Then since M and $M^\perp$ are subspaces, we have

$$x + y = \underbrace{(x_M + y_M)}_{\in M} + \underbrace{(x_{M^\perp} + y_{M^\perp})}_{\in M^\perp}.$$

Then for any $\alpha, \beta \in \mathbb{R}$ we have

$$P(\alpha x + \beta y) = \alpha x_M + \beta y_M = \alpha P(x) + \beta P(y).$$

So P is linear.

For any $x \in H$,

$$\|x\| = \|x_M + x_{M^\perp}\| = \sqrt{\langle x_M + x_{M^\perp} | x_M + x_{M^\perp} \rangle} = \sqrt{\langle x_M | x_M \rangle + \langle x_{M^\perp} | x_{M^\perp} \rangle} = \sqrt{\|x_M\|^2 + \|x_{M^\perp}\|^2}.$$

Writing down the expression of the bound of the operator P (if it exists)

$$\|p\| = \sup_x \frac{\|P(x)\|}{\|x\|} = \sup_x \frac{\|x_M\|}{\sqrt{\|x_M\|^2 + \|x_{M^\perp}\|^2}} \leq 1.$$

Obviously, if $m \in M$, we have $\|P(m)\| = \|m\|$; since M is more than merely the origin thee exists $m \in M, m \neq 0$ and we see that the bound of 1 is achieved. This is where the assumption of M being of at least one-dimension comes into play. □