ECE 555 **Assignment # 4**

Issued: February 24 Due: March 10, 2011

Reading: Continue reading CTCN: We will cover Sections 9.1–9.5.

Exercises:

10. Consider the linear model $X(t+1) = AX(t) + BU(t) + N(t+1)$, with $\boldsymbol{N}$ i.i.d., zero mean, and finite variance. Consider the quadratic cost,

$$c(x,u) = \tfrac{1}{2}x'Qx + \tfrac{1}{2}u'Ru$$

where $R > 0$ and $Q \geq 0$. Apply one step of the VIA algorithm to compute V_1 *by hand* when V_0 is a quadratic. *Easy, right*?
Explain why computation is it so much harder for the ℓ_1 control problem, with $c(x,u) = \sqrt{x'Qx} + \sqrt{u'Ru}$.

Solution First we respond to " *Easy, right*": With quadratic cost, computation of V_1 is obtained in closed form, and V_1 is quadtric. The ℓ_1 problem is "*easy*" because it is a convex program, but it is hard because we cannot obtain a closed-form expression V_1.

Assume $V_0 = \frac{1}{2}x'Mx$ where X is a *positive definite* matrix. We now compute V_1:

$$\begin{aligned}
V_1(x) &= \min_u \big(c(x,u) + P_u V_0\big) \\
&= \min_u \big(c(x,u) + \mathsf{E}[\tfrac{1}{2}(Ax+Bu+N(1))'M(Ax+Bu+N(1))]\big) \\
&= \min_u \big(c(x,u) + \mathsf{E}[\tfrac{1}{2}u'B'MBu + x'A'MBu + N(1)'M(Ax+Bu)] \\
&\quad +\mathsf{E}[\tfrac{1}{2}x'A'MAx + \tfrac{1}{2}N(1)'AN(1)]\big) \\
&= \min_u \big(c(x,u) + \tfrac{1}{2}u'B'MBu + x'A'MBu \\
&\quad +\mathsf{E}[N(1)'M(Ax+Bu)] + \tfrac{1}{2}x'A'MAx + \tfrac{1}{2}\mathsf{E}[N(1)'AN(1)]\big) \\
&= \min_u \big(\tfrac{1}{2}u'Ru + \tfrac{1}{2}u'B'MBu + x'A'MBu\big) \\
&\quad +\tfrac{1}{2}x'Qx + \tfrac{1}{2}x'A'MAx + \tfrac{1}{2}\mathsf{E}[N(1)'AN(1)] \qquad (1)
\end{aligned}$$

To find the minimizer u, we use the necessary condition for optimality:

$$\begin{aligned}
0 &= \nabla_u\big(\tfrac{1}{2}u'Ru + \tfrac{1}{2}u'B'MBu + x'A'MBu\big) \\
&= Ru + B'MBu + B'MAx
\end{aligned}$$

Thus, we have

$$u = -(R + B'MB)^{-1}B'MAx$$

Substitute this into (1), we obtain

$$V_1 = \tfrac{1}{2}x'(Q + A'MA)x - \tfrac{1}{2}x'A'MB(R+B'MB)^{-1}B'MAx + \tfrac{1}{2}\mathsf{E}[N(1)'AN(1)]$$

Applying matrix inversion lemma, we obtain

$$(M^{-1} + BR^{-1}B')^{-1} = M - MB(R + B'MB)^{-1}B'M$$

Consequently, we have

$$V_1 = \tfrac{1}{2}x'\big(Q + (M^{-1} + BR^{-1}B')^{-1}\big)x + \tfrac{1}{2}\mathsf{E}[N(1)'AN(1)]$$

where $Q + (M^{-1} + BR^{-1}B')^{-1}$ is clearly a positive definite matrix.

11. Recall the controlled queueing model,

$$Q(t+1) = Q(t) - U(t) + A(t+1) \tag{2}$$

in which $\boldsymbol{A}$ is i.i.d., and $U(t), Q(t)$ are non-negative valued. We have considered the cost function $c(x,u) = x + \frac{1}{2}u^2$, and found that the relative value function can be approximated by the fluid value function:

$$\min_u \bigl\{ c(x,u) + D_u h\,(x) \bigr\} \approx \eta$$

where $h(x) = kx^p$, with $p = 3/2$, and k, η are constants.
See if you can find an approximation for the discounted-cost optimal control problem. If desired, you can take the form given in CTCN: Find a function h such that,

$$\min_u \bigl\{ c(x,u) + D_u h\,(x) \bigr\} \approx \gamma h(x)$$

where $\gamma > 0$ (corresponding to discount factor $\beta = (1+\gamma)^{-1}$). This simplifies comparison with the fluid model total-cost problem.

Solution Consider the discounted cost optimality equation (DCOE) for the fluid model,

$$\min_u \bigl\{ c(x,u) + D_u J\,(x) \bigr\} = \gamma J(x) \tag{3}$$

in which $D_u J\,(x) = (-u+\alpha)\cdot \nabla J\,(x)$. The DCOE is solved with the fluid value function,

$$J^*(x) = \inf_{u(\,\cdot\,)} \int_0^\infty e^{-\gamma t} c(q(t), u(t))\, dt, \qquad q(0) = x \ge 0.$$

To simplify the problem we take $\alpha = 0$, so the fluid model is defined by $\dot q = -u$. In this case, when $\gamma = 0$, we can obtain an explicit solution,

$$J_0^*(x) = kx^{3/2}\,, \qquad \text{where } k = 4\sqrt{2}/3. \tag{4}$$

For $\gamma > 0$ we can apply (3) and the form of c to deduce $u^* = -\frac{d}{dx}J^*(x)$, and

$$x - \tfrac{1}{2}\bigl(\tfrac{d}{dx}J^*(x)\bigr)^2 = \gamma J^*(x) \tag{5}$$

We can't solve this, but note that for any non-negative, smooth function $J\colon \mathbb{R}_+ \to \mathbb{R}_+$, on setting $c_J(x,u) = \gamma J(x) + \frac{1}{2}(\frac{d}{dx}J(x))^2 + \frac{1}{2}u^2$, we solve the DCOE using this function,

$$\min_u \bigl\{ c_J(x,u) + D_u J\,(x) \bigr\} = \gamma J(x) + \tfrac{1}{2}(\tfrac{d}{dx}J(x))^2 - \tfrac{1}{2}\bigl(\tfrac{d}{dx}J(x)\bigr)^2 = \gamma J(x) \tag{6}$$

Our goal is to find J so that $c_J \approx c$.

We first establish some structure of the solution to (5):

(i) Convexity of c and linearity of the fluid model implies that J^* is convex.

(ii) If $x \sim 0$ or $\gamma \sim 0$ then the exponential $e^{-\gamma t}$ has little impact. Hence $J^*(x) = O(x^{3/2})$ for $x \sim 0$ (the structure of J^* when $\gamma = 0$).

(iii) An upper bound on J^* is obtained on setting $u(t) = u_0 1\{q(t) > 0\}$. The resulting value function has linear growth, which implies that $J^*(x) = O(x)$ for large x. This combined with convexity implies that the following limit exists, and is finite:

$$\tfrac{d}{dx} J^*(\infty) := \lim_{x\to\infty} \tfrac{d}{dx} J^*(x) = \lim_{x\to\infty} \frac{1}{x} J^*(x) < \infty$$

A second look at the DCOE gives the value, $\tfrac{d}{dx} J^*(\infty) = \gamma^{-1}$.

Putting this all together, the solution to (5) can be approximated by a convex function with linear growth. There are many such approximations.

If we want $J(x) = O(x^{3/2})$ for $x \sim 0$, then we could try

$$J(x) = \frac{x}{\gamma}(1 - e^{-k\gamma\sqrt{x}})$$

This has the right shape for large x, and is approximately $kx^{3/2}$ when $x \sim 0$. It is convex, with

$$\tfrac{d}{dx} J(x) = \frac{1}{\gamma}\Big((1 - e^{-k\gamma\sqrt{x}}) + \tfrac{1}{2}k\gamma\sqrt{x}e^{-k\gamma\sqrt{x}}\Big)$$

Consequently, $c_J(x) \sim x + u^2$ (in engineering terms!) for arbitrary γ.
We have $J(x) \to kx^{3/2}$ $\gamma \downarrow 0$, for each $x \ge 0$. Considering (4), a natural choice for k is $k = 4\sqrt{2}/3$. In this case it follows that $c_J(x) \to x + u^2$ as well.

12. Returning once more to (2), in the average cost setting, note that the cost function $c(x, u) = x + \frac{1}{2}u^2$ is not well motivated — *why sum the two costs*?
Let's consider instead the constrained optimization problem,

$$\textbf{min} \quad \mathsf{E}[U(\infty)^2] \qquad \textbf{s.t.} \quad \mathsf{E}[X(\infty)] \le \bar{\eta} \tag{7}$$

where $\bar{\eta}$ is a pre-specified constraint.
Approximate the solution to this problem, with $\bar{\eta}$ half the steady-state cost obtained when $U(t) = 1\{X(t) \ge 1\}$ (see Theorem 3.0.1 in CTCN).
To solve this problem you must truncate the state space, and you should assume that $(U(k), Q(k), A(k))$ are restricted to an integer lattice. I'll give you some flexibility in modeling: the marginal of $\boldsymbol{A}$ has mean near 10, and variance between 5 and 25 — for example, you can choose a uniform, or geometric distribution.
You can solve an LP, or you can compute the solution to the average cost optimization problem with $c(x, u) = \lambda x + \frac{1}{2}u^2$, for various $\lambda > 0$.

Solution Regardless of what approach you take, the state space and action space must be truncated: $\mathsf{X} = [0, \dots N]$ and $\mathsf{U} = [0, \dots M]$ for some finite N, M. We also have state dependent input constraints:

$$\mathsf{U}(x) = \{x : u \in \mathsf{U},\ u \le x\}$$

There are two ways to approach this problem:

Method 1: Lagrangian relaxation For each $\lambda \ge 0$ solve the MDP

$$\textbf{min} \quad \mathsf{E}[U(\infty)^2] + \lambda\big(\mathsf{E}[X(\infty)] - \bar{\eta}\big)$$

Find the value $\lambda^* > 0$ such that the resulting policy gives $\mathsf{E}[X^*(\infty)] = \bar{\eta}$, and this will solve the constrained optimization problem (7).

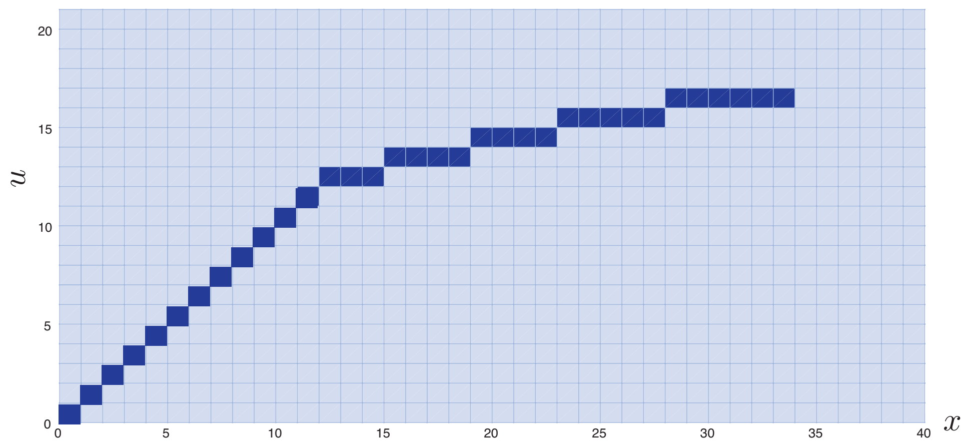
Method 2: LP The constrained optimization problem can be written as the following linear programming problem:

$$\begin{array}{ll} \max & \sum_{x,u} \Gamma(x,u) c_{\text{U}}(u) \\ \text{s.t.} & \sum_{x,u} \Gamma(x,u) c_{\text{X}}(x) \le \bar{\eta} \\ & \sum_{x,u} \Gamma(x,u) = 1 \\ & \sum_{x,u} \Gamma(x,u) P_u(x,y) = \sum_u \Gamma(y,u) \quad \forall y \\ & \Gamma(x,u) \ge 0 \quad \forall x,u \end{array} \tag{8}$$

where $c_{\text{U}}(u) \equiv u^2$, $c_{\text{X}}(x) \equiv x$, and where $P_u(x,y)$ is the transition law under control u for this Markov model,

$$Q(t+1) = [Q(t) - U(t) + A(t+1)]_0^N$$

Here is the result you would expect for Γ^*,



where the dark color denotes the support of Γ. In this particular experiment, the solution obtained was deterministic. In other experiments, there will be a few values of x for which $\Gamma(x,u) \in (0,1)$, so that the policy is randomized.

In the case illustrated in the figure, the optimizing Γ^* defines a (deterministic) state feedback policy ϕ^*. For small x, we have $\phi^*(x) = x$ (the queue is emptied out). For larger x, the policy has the $\sqrt{x}$-shape predicted by the analysis of the fluid model.