B.3 Problem Set III

(First Part due Wed. Oct. 21)[3]

Problem B.3.1. Project Proposal Due Wed. Oct. 21 (~ 1 page). Reminders about the project: pick a topic of your interest in dynamic programming. You can work on your own or in pairs. I encourage you to talk to me before so that we can agree on the proposal. Ideally your project would include some personal input, an application of DP to your research, etc. Otherwise you can also report on say $2-3$ paper on a topic related to the class. In this case I do not want a superficial summary of the papers, but a deep understanding and a critical evaluation of the papers.

Administrative stuff: we will use a blog to manage feedback regarding the projects. The address is
`http://cooperativecontrol.wordpress.com/`
I've posted already a bit more administrative information there. Please create an account on `WordPress.com` by Wed. October 14. Then send me the email address that you used to create that account, so that I can add you as a contributor to the blog.

Problem B.3.2. Watch Stephen Boyd's lectures on MPC, available at
`http://www.stanford.edu/class/ee364b/videos.html`.
See also the link reference section on the course web page. The lectures are the ones of May 22 (Model predictive control) and May 27 (Stochastic model predictive control).

Problem B.3.3. Do the exercise in chapter 11 (on the performance of one-step lookahead policies; I gave the proof in class, but it's useful to do it again).

Problem B.3.4. Consider the following two-stage example, due to [TW66], which involves the following two-dimensional linear system with scalar control and disturbance:

$$x_{k+1} = x_k + bu_k + dw_k, \ k = 0, 1,$$

where $b = [1, 0]^T$ and $d = [1/2, \sqrt{2}/2]^T$. The initial state is $x_0 = 0$. The controls u_0 and u_1 are unconstrained. The disturbances w_0 and w_1 are independent random variables and each takes the values 1 and -1 with equal probability 1/2. Perfect information prevails. The cost is

$$E_{w_0,w_1}[\|x_2\|],$$

where $\|\cdot\|$ denotes the usual Euclidian norm. Show that the CEC with nominal values $\bar{w}_0 = \bar{w}_1 = 0$ has worse performance than the optimal open-loop controller. In particular, show that the optimal open-loop cost and the optimal closed-loop cost are both $\sqrt{3}/2$, but the cost of corresponding to CEC is 1.

[3] this version: Oct. 6 2009

Problem B.3.5 (Continuous Space Shortest Path Problems)**.** Consider the two-dimensional system

$$\frac{d}{dt}x_1 = u_1, \quad \frac{d}{dt}x_2 = u_2,$$

with the control constraint $\|u(t)\| = 1$. We want to find a state trajectory that starts at a given point $x(0)$, ends at another point $x(T)$, and minimizes

$$\int_0^T r(x(t))dt.$$

The function $r(\cdot)$ is nonnegative and continuous, and the final time T is subject to optimization. Suppose we discretize the plane with a mesh of size Δ that passes through $x(0)$ and $x(T)$, and we introduce a shortest path problem of going from $x(0)$ to $x(T)$ using moves of the following type: from each mesh point $\bar{x} = (\bar{x}_1, \bar{x}_2)$, we can go to each of the mesh points $(\bar{x}_1 + \Delta, \bar{x}_2)$, $(\bar{x}_1 - \Delta, \bar{x}_2)$, $(\bar{x}_1, \bar{x}_2 + \Delta)$ and $(\bar{x}_1, \bar{x}_2 - \Delta)$, at a cost $r(\bar{x})\Delta$. Show by an example that this is a bad discretization of the original problem in the sense that the shortest distance need not approach the optimal cost of the original problem as $\Delta \to 0$.

Problem B.3.6 (Convergence Properties of Rank-One Correction)**.** Consider the solution of the system $J = FJ$, where $F : \mathbb{R}^n \to \mathbb{R}^n$ is the mapping

$$FJ = h + QJ,$$

h is a given vector in $\mathbb{R}^n$, and Q is an $n \times n$ matrix. Consider the generic rank-one correction iteration $J := MJ$, where $M : \mathbb{R}^n \to \mathbb{R}^n$ is the mapping

$$MJ = FJ + \gamma z,$$

and

$$z = Qd, \quad \gamma = \frac{(d-z)^T(FJ - J)}{\|d - z\|^2}.$$

1. Show that any solution J^* of the system $J = FJ$ satisfies $J^* = MJ^*$.

2. Verify that the value iteration method that uses the error bounds in the manner of

$$\hat{J}_k = T^k J + \left[\frac{\alpha}{n(1-\alpha)} \sum_{i=1}^{n} ((T^k J)(i) - (T^{k-1}J)(i)) \right] e$$

   (see Eq. (1.21) in Bertsekas' book volume II) is a special case of the iteration $J := MJ$ with d equal to the all-ones vector $e = [1, \ldots, 1]^T$.

3. Assume that d is an eigenvector of Q, let λ be the corresponding eigenvalue, and let $\lambda_1, \ldots, \lambda_{n-1}$ be the remaining eigenvalues. Show that MJ can be written as

$$MJ = \tilde{h} + RJ$$

where $\tilde{h}$ is some vector in $\mathbb{R}^n$ and

$$R = Q - \frac{\lambda}{(1-\lambda)\|d\|^2} dd^T (I - Q).$$

Show also that $Rd = 0$ and that for all k and J,

$$R^k = RQ^{k-1}, \ \ M^k J = M(F^{k-1} J).$$

Furthermore, the eigenvalues of R are $0, \lambda_1, \ldots, \lambda_{n-1}$ (this last statement requires a somewhat complicated proof).

4. Let d be as in part 3, and suppose that $e_1, \ldots, e_{n-1}$ are eigenvectors corresponding to $\lambda_1, \ldots, \lambda_{n-1}$. Suppose that a vector J can be written as

$$J = J^* + \xi e + \sum_{i=1}^{n-1} \xi_i e_i,$$

where J^* is a solution of the system. Show that, for all $k > 1$,

$$M^k J = J^* + \sum_{i=1}^{n-1} \xi_i \lambda_i^{k-1} R e_i,$$

so that if λ is a dominant eigenvalue and $\lambda_1, \ldots, \lambda_{n-1}$ lie within the unit circle, $M^k J$ converges to J^* at a rate governed by the subdominant eigenvalue.

Due Monday November 2: For the next two computational problems, you are encouraged to work in groups of $2-3$ people. You can submit a single report per group.

Problem B.3.7. Solve problem B.1.7 using an MPC controller. For the deterministic case, experiment with the length of the horizon and compare the cost obtained with the optimal cost you computed in problem set 1. For the problem with noise, use a certainty equivalent model predictive controller for the stochastic problem (this is what is described in the second video mentioned above). Again experiment with the length of the horizon and report your findings, provide some plots of the trajectories obtained, and discuss your design by comparing it to the designs you obtained in the previous problem sets. As before, consider a full-state feedback controller (i.e., of the form $\mu(x)$).

Problem B.3.8. Consider the inventory control problem with backlogs, as treated in pp. 21-22 of Volume I of the text. [The DP equation (1.4) given there is a bit more convenient than its counterpart (4.21) on p. 162.] Let

$$c = 1, r(x) = p \max(0, -x) + h \max(0, x),$$

with $p = 4, h = 2, N = 10$.

1. Let the demands w_k be i.i.d., with a discrete uniform distribution in the set $\{0, 1, ..., 10\}$. Calculate and plot the function J_0, and find an optimal policy.

2. Let now the demands be i.i.d. with a continuous uniform distribution on the set $[0, 10]$. Calculate and plot the function J_0, and find an optimal policy. (Note: this cannot be done exactly. You may either approximate the problem by a discrete one and solve the discrete problem exactly, or work with the DP equations for the continuous problem and approximate when needed, e.g., replace integrals by sums. In either case, explain what you did and defend the number of discrete grid points that you used. (You may Þnd reading Section 6.6.1 to be useful.)

3. Same setup as in part 2, but we will now use an *approximation architecture for the cost-to-go function*. The main idea is to calculate values for the cost-to-go at a finite set of state-time pairs, then to make a "least squares" fit of these values with a function of a given type, such as a polynomial function (see Section 6.4.3, where the same idea is used to approximate the cost-to-go of the base policy). We first explain this methodology. In particular, suppose we have calculated the correct value of the optimal cost-to-go $J_{N-1}(x^i)$ at the next to last stage stage for m states $x^1, \ldots, x^m$ through the DP formula

$$J_{N-1}(x) = \min_u E_w[g(x, u, w) + J_N(f(x, u, w))],$$

and the given terminal cost function J_N (for simplicity, we drop the subscripts of g, x, u, w in the following, but this is not necessary for this methodology). We can then approximate the entire $J_{N-1}(x)$ by a function of the form $\tilde{J}_{N-1}(x; r_{N-1})$, where r_{N-1} is a vector of parameters which can be obtained by solving the least squares problem

$$\min_r \sum_{i=1}^{m} \left| J_{N-1}(x^i) - \tilde{J}_{N-1}(x^i; r) \right|^2.$$

For example, if $x \in \mathbb{R}^n$ and $\tilde{J}_{N-1}$ is specified to be linear, the vector r consists of $\alpha \in \mathbb{R}^n$ and scalar β and $\tilde{J}_{N-1}(x, r) = \alpha^T x + \beta$. The least squares problem then is

$$\min_{\alpha,\beta} \sum_{i=1}^{m} \left| J_{N-1}(x^i) - \alpha^T x^i - \beta \right|^2.$$

Once an approximating function $\tilde{J}_{N-1}(x; r_{N-1})$ is obtained, it can be used to obtain an approximating function $\tilde{J}_{N-2}(x; r_{N-2})$. In particular, (approximate) cost-to-go function values $\hat{J}_{N-2}(x^i)$ are obtained for m states $x^1, \ldots, x^m$ through the (approximate) DP formula

$$\hat{J}_{N-2}(x) = \min_u E_w[g(x, u, w) + \tilde{J}_{N-1}(f(x, u, w); r_{N-1})].$$

These values are then used to approximate the cost-to-go function $J_{N-2}(x)$ by a function of the form $\tilde{J}_{N-2}(x; r_{N-2})$, by solving the problem

$$\min_r \sum_{i=1}^{m} \left| \hat{J}_{N-2}(x^i) - \tilde{J}_{N-2}(x^i; r) \right|^2 .$$

The process can be continued to obtain $\tilde{J}_k(x; r_k)$ up to k = 0. Given approximate cost-to-go functions $\tilde{J}_0(x; r_0), \ldots, \tilde{J}_{N-1}(x; r_{N-1})$, one may obtain a suboptimal policy by using at state-time pair (x, k) the control

$$\tilde{\mu}_k(x) = \arg\min_u \left\{ E_w[g(x, u, w) + \tilde{J}_{k+1}(f(x, u, w); r_{k+1})] \right\} .$$

We now return to part 2 of the problem, where value functions are to be approximated by quadratics, of the form $ax^2 + bx + c$. Note that with this approximation architecture, the expected values that are needed in the DP algorithm can be evaluated analytically. Compare the results with those obtained in part 2.