

MECH 6326 - Optimal Control and Dynamic Programming

Homework 1

Jonas Wagner
jonas.wagner@utdallas.edu

2023, February 10st

Contents

Problem 1	2
Problem 2	4
a)	4
b)	4
c)	5
Problem 3	8
a)	8
b)	8

Problem 1

PROBLEM:

Find and research an example (ideally from your own research experience/interests, or from an actual company/startup) of a problem of multi-step decision making under uncertainty in the real world. Write a short narrative that briefly describes i) the use (or potential use) of closed-loop feedback control/recourse due to uncertain factors affecting performance and operation and ii) the performance objectives and operation of the system, including the system dynamics, system states, inputs, outputs, and stochastic disturbances and other uncertainties. Comment specifically on any stochastic uncertainties and how you might go about collecting data to describe their probability distribution. Explain why an open-loop controller may not be sufficient and what are the advantages of using closed-loop feedback/recourse by utilizing knowledge of the system state and/or realizations of uncertain parameters at each time instant

SOLUTION:

Since I have many examples that could fit this description I will take one of the more complicated ones to touch base on each of the questions.

NOVA Vehicle [1] Specifically I have a bicycle model of the vehicle that I've been using in research. I can copy over my entire report from Robust Control on the various uncertainties and how we model this (in continuous time, I have been working with a DT model for a lateral motion in research) but I'll instead just have the link here: https://github.com/jonaswagner2826/MECH6323/blob/master/FinalProject/Report/MECH6323_FinalProject.pdf

The TLDR for the system is included with the Figure1 and the nonlinear dynamics in (1).

As for the specific questions asked, here's a bit of a narrative:

(i) For 'simpler' vehicle motions (lane-keeping) we can split the problem into partially independent longitudinal ('Speed/Position') and lateral ('Side-to-side/steering') controllers. Each of these can benefit from a closed-loop controller that accounts for many areas of system and sensor noises.

(ii)

$$\begin{cases} I\ddot{\theta} &= aF_{\eta f}\delta + bF_{\zeta f} - bF_{\zeta r} \\ m(\dot{V}_{\zeta} + V_{\eta}\dot{\theta}) &= F_{\eta f}\delta + F_{\zeta f} + F_{\zeta r} \\ m(\dot{V}_{\eta} + V_{\zeta}\dot{\theta}) &= F_{\eta f} + F_{\eta r} + F_{\zeta f}\delta \\ \dot{x} &= -V_{\zeta}\sin(\theta) + V_{\eta}\cos(\theta) \\ \dot{y} &= V_{\eta}\cos(\theta) + V_{\zeta}\sin(\theta) \end{cases} \quad (1)$$

The global system coordinates are (x, y, θ) , longitudinal (η) and lateral (ζ) velocities (V_{η}, V_{ζ}), longitudinal and lateral forces for the front and rear as $(F_{\eta f}, F_{\eta r}, F_{\zeta f}, F_{\zeta r})$, steering angle of δ , and with physical parameters that are uncertain.

The following is the implementation of the nonlinear dynamics equations (1) in state space with the state

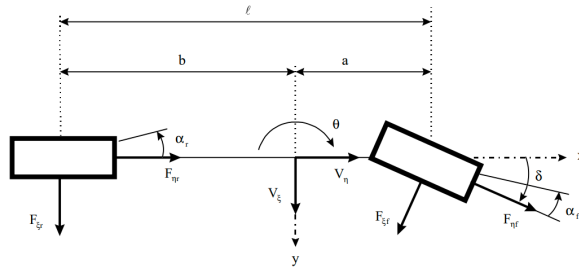


Figure 1: 3 DOF Bicycle Model

vector, $x = [\dot{\theta}, V_\zeta, V_\eta, x, y, \theta]^T$, and input vector $u = [\delta, F_{\eta f}, F_{\eta r}, F_{\zeta f}, F_{\zeta r}]^T$.

$$\begin{cases} \dot{x}_1 &= \frac{aF_{\eta,f}\delta + bF_{\zeta,f} - bF_{\zeta,r}}{I} \\ \dot{x}_2 &= \frac{F_{\eta,f}\delta + F_{\zeta,f} - F_{\zeta,r}}{m} - x_3x_1 \\ \dot{x}_3 &= \frac{F_{\eta,f} + F_{\eta,r} - F_{\zeta,r}\delta}{m} - x_2x_1 \\ \dot{x}_4 &= -x_2 \sin(x_6) + x_3 \cos(x_6) \\ \dot{x}_5 &= x_2 \cos(x_6) + x_3 \sin(x_6) \\ \dot{x}_6 &= x_1 \end{cases} \quad (2)$$

In addition to any model uncertainty (which can be modeled as system noise) the stochastic uncertainty is generated from sensor measurements and nonlinearities of actuators. See the report for all the specifics on the sensors and actuators which have explanations on the uncertainty for each.

The advantages of the of a closed-loop operation (and additionally the dynamic updating of parameters) is a very long list so I'm going to neglect explaining them, but it essentially is a straight-forward answer of: uncertainty makes it act different then expected so the open-loop will likely fail.

Problem 2

Consider the DT L(TI) stochastic dynamical system

$$x_{k+1} = Ax_k + Bu_k + w_k \quad (3)$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, and $w_k \sim \mathcal{N}(0, W)$ is an iid signal. Suppose $x_0 \sim \mathcal{N}(0, X_0)$ is another iid random variable. The control input is given by a linear feedback control policy $u_k = Kx_k$, with $K \in \mathbb{R}^{m \times n}$.

a)

PROBLEM:

Let the state mean be defined as $\mu_k := E[x_k]$ and state second moment matrix defined as $\Sigma_k := E[x_k x_k^T]$. Derive the dynamics for the state mean and second moment matrix, (i.e.) write μ_{k+1} and Σ_{k+1} in terms of μ_k , Σ_k and the given problem data (A, B, K, W, X_0) .

SOLUTION:

Since the system is LTI, the propagation of the mean and variance can be directly computed as the expected value of the update equations.

In the closed loop we have

$$x_{k+1} = (A + BK)x_k + w_k$$

State mean update:

$$\begin{aligned} \mu_{k+1} &= E[x_{k+1}] = E[(A + BK)x_k + w_k] \\ &= E[(A + BK)x_k] + E[w_k] \\ &= (A + BK)E[x_k] + (0) \\ &= (A + BK)\mu_k \end{aligned}$$

State covariance update:

$$\begin{aligned} \Sigma_{k+1} &= E[x_{k+1} x_{k+1}^T] = E[((A + BK)x_k + w_k)((A + BK)x_k + w_k)^T] \\ &= E[((A + BK)x_k + w_k)(x_k^T (A + BK)^T + w_k^T)] \\ &= E[(A + BK)x_k x_k^T (A + BK)^T + (A + BK)x_k w_k + w_k x_k^T (A + BK)^T + w_k w_k^T] \\ &= E[(A + BK)x_k x_k^T (A + BK)^T] + E[(A + BK)x_k w_k] + E[w_k x_k^T (A + BK)^T] + E[w_k w_k^T] \\ &= (A + BK)E[x_k x_k^T](A + BK)^T + (A + BK)E[x_k w_k] + E[w_k x_k^T](A + BK)^T + E[w_k w_k^T] \end{aligned}$$

since x_k and w_k are independent,

$$\begin{aligned} &= (A + BK)E[x_k x_k^T](A + BK)^T + (A + BK)E[x_k]E[w_k] + E[w_k]E[x_k^T](A + BK)^T + E[w_k w_k^T] \\ &= (A + BK)(\Sigma_k)(A + BK)^T + (A + BK)(\mu_k)(0) + (0)(\mu_k)(A + BK)^T + (W) \\ &= (A + BK)\Sigma_k(A + BK)^T + W \end{aligned}$$

b)

PROBLEM:

Suppose $n = m = 1$. Derive an expression for the steady state second moment $\Sigma_\infty := \lim_{k \rightarrow \infty} \Sigma_k$. Under what conditions does this converge to a finite value? In this case is $\lim_{k \rightarrow \infty} \mathbf{P}(x_k \geq 3\sqrt{\Sigma_\infty})$; (i.e.) the probability that in steady state, the state takes a value greater than $3\sqrt{\Sigma_\infty}$?

SOLUTION:

Steady state occurs when $\Sigma_{k+1} = \Sigma_k = \Sigma_\infty$.

$$\Sigma_\infty = (A + BK)\Sigma_\infty(A + BK)^T + W$$

since $(A + BK)$ is a scalar,

$$\begin{aligned}\Sigma_\infty &= (A + BK)^2\Sigma_\infty + W \\ (1 - (A + BK)^2)\Sigma_\infty &= W \\ \Sigma_\infty &= \frac{W}{1 - (A + BK)^2}\end{aligned}$$

Σ_k will always converge when $A + BK \neq 1$. Assuming that K is designed such that $A + BK$ is stable, Σ_k will always be finite.

Since μ_k should be driven to 0, $\mathbf{P}(x_k \geq 3\sqrt{\Sigma_\infty}) = \mathbf{P}(z \geq 3) = 0.13\%$.

c)

PROBLEM:

Instead of a linear feedback control policy, derive an expression for an *open-loop control input sequence* $\mathbf{u} = [u_0^T, u_1^T, \dots, u_{N-1}^T]^T$ that optimizes the expected finite-horizon cost for a fixed initial state x_0 with cost matrices $G \succ 0$ and $R \succ 0$.

$$E \left[\sum_{k=1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_T^T Q x_T \right] \quad (4)$$

SOLUTION:

[2] Let $\mathbf{x} = [x_1^T, x_1^T, \dots, x_T^T]^T$, $\mathbf{u} = [u_0^T, u_1^T, \dots, u_{T-1}^T]^T$, and $\mathbf{w} = [w_0^T, w_1^T, \dots, w_{T-1}^T]^T$. \mathbf{x} can then be redefined in terms of \mathbf{u} , \mathbf{w} , and x_0 :

Since

$$x_{k+1} = Ax_k + Bu_k$$

we have

$$\begin{aligned}x_1 &= Ax_0 + Bu_0 + w_0 \\ x_2 &= Ax_1 + Bu_1 + w_1 \\ &= A(Ax_0 + Bu_0 + w_0) + Bu_1 + w_1 \\ &= A^2x_0 + ABu_0 + Bu_1 + Aw_0 + w_1 \\ x_3 &= A^3x_0 + A^2Bu_0 + ABu_1 + Bu_2 + A^2w_0 + Aw_1 + w_2 \\ &\vdots \\ x_N &= Ax_{N-1} + Bu_{N-1} \\ &= A^N x_0 + A^{N-1}Bu_0 + \dots + ABu_{N-1} + Bu_N + w_N\end{aligned}$$

Additionally, this can be generalized as:

$$x_t = A^t x_0 + \sum_{i=0}^{t-1} A^{t-i} (B\mathbf{u}_i + \mathbf{w}_i)$$

This can be written in vector form as

$$\mathbf{x} = Ax_0 + B\mathbf{u} + C\mathbf{w} \quad (5)$$

where

$$\mathbf{A} = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} B & 0 & \dots & 0 \\ AB & B & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ A^{N-1}B & \dots & AB & B \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} I & 0 & \dots & 0 \\ A & I & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ A^{N-1} & \dots & A & I \end{bmatrix}$$

The LQR problem defined in (4) is equivalent to

$$J = E \left[\sum_{k=1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_T^T Q x_T \right] = E[\mathbf{x}^T Q \mathbf{x} + \mathbf{u}^T R \mathbf{u}]$$

and thus,

$$\begin{aligned} J &= E[\mathbf{x}^T Q \mathbf{x} + \mathbf{u}^T R \mathbf{u}] \\ &= E[\mathbf{x}^T Q \mathbf{x}] + E[\mathbf{u}^T R \mathbf{u}] \\ &= E[(Ax_0 + B\mathbf{u} + C\mathbf{w})^T Q (Ax_0 + B\mathbf{u} + C\mathbf{w})] + \mathbf{u}^T R \mathbf{u} \\ &= E[(x_0^T A^T + \mathbf{u}^T B^T + \mathbf{w}^T C^T) Q (Ax_0 + B\mathbf{u} + C\mathbf{w})] + \mathbf{u}^T R \mathbf{u} \\ &= E[x_0^T A^T Q Ax_0 + x_0^T A^T Q B \mathbf{u} + x_0^T A^T Q C \mathbf{w} \\ &\quad + \mathbf{u}^T B^T Q Ax_0 + \mathbf{u}^T B^T Q B \mathbf{u} + \mathbf{u}_T^T B^T Q C \mathbf{w} \\ &\quad + \mathbf{w}^T C^T Q Ax_0 + \mathbf{w}^T C^T Q B \mathbf{u} + \mathbf{w}^T C^T Q C \mathbf{w}] \\ &\quad + \mathbf{u}^T R \mathbf{u} \\ &= E[x_0^T A^T Q Ax_0] + E[x_0^T A^T Q B \mathbf{u}] + E[x_0^T A^T Q C \mathbf{w}] \\ &\quad + E[\mathbf{u}^T B^T Q Ax_0] + E[\mathbf{u}^T B^T Q B \mathbf{u}] + E[\mathbf{u}_T^T B^T Q C \mathbf{w}] \\ &\quad + E[\mathbf{w}^T C^T Q Ax_0] + E[\mathbf{w}^T C^T Q B \mathbf{u}] + E[\mathbf{w}^T C^T Q C \mathbf{w}] \\ &\quad + \mathbf{u}^T R \mathbf{u} \end{aligned}$$

since u is not stochastic and both x_0 and \mathbf{w} are linearly independent,

$$\begin{aligned} &= E[x_0^T A^T Q Ax_0] + E[x_0^T A^T Q B \mathbf{u}] + E[x_0^T A^T Q C E[\mathbf{w}]] \\ &\quad + \mathbf{u}^T B^T Q A E[x_0] + \mathbf{u}^T B^T Q B \mathbf{u} + \mathbf{u}_T^T B^T Q C E[\mathbf{w}] \\ &\quad + E[\mathbf{w}^T] C^T Q A E[x_0] + E[\mathbf{w}^T] C^T Q B \mathbf{u} + E[\mathbf{w}^T C^T Q C \mathbf{w}] \\ &\quad + \mathbf{u}^T R \mathbf{u} \\ &= E[x_0^T A^T Q Ax_0] + (\mu_0) A^T Q B \mathbf{u} + (\mu_0) A^T Q C (0) \\ &\quad + \mathbf{u}^T B^T Q A (\mu_0) + \mathbf{u}^T B^T Q B \mathbf{u} + \mathbf{u}_T^T B^T Q C (0) \\ &\quad + (0) C^T Q A (\mu_0) + (0) C^T Q B \mathbf{u} + E[\mathbf{w}^T C^T Q C \mathbf{w}] \\ &\quad + \mathbf{u}^T R \mathbf{u} \\ &= E[x_0^T A^T Q Ax_0] + \mu_0 A^T Q B \mathbf{u} + \mathbf{u}^T B^T Q A \mu_0 \\ &\quad + \mathbf{u}^T B^T Q B \mathbf{u} + E[\mathbf{w}^T C^T Q C \mathbf{w}] + \mathbf{u}^T R \mathbf{u} \\ J &= \mathbf{u}^T (B^T Q B + R) \mathbf{u} + \mu_0 A^T Q B \mathbf{u} + \mathbf{u}^T B^T Q A \mu_0 + E[\mathbf{w}^T C^T Q C \mathbf{w}] + E[x_0^T A^T Q Ax_0] \end{aligned}$$

The optimal open loop control policy can then be found by minimizing the LQR problem:

$$\min_{\mathbf{u}} J = \mathbf{u}^T (B^T Q B + R) \mathbf{u} + \mu_0 A^T Q B \mathbf{u} + \mathbf{u}^T B^T Q A \mu_0 + E[\mathbf{w}^T C^T Q C \mathbf{w}] + E[x_0^T A^T Q Ax_0]$$

We can find the optimal solution by setting the gradient of the objective function to zero:

$$\begin{aligned} 0 = \nabla_u J &= \nabla_u (\mathbf{u}^T (B^T Q B + R) \mathbf{u} + \mu_0 A^T Q B \mathbf{u} + \mathbf{u}^T B^T Q A \mu_0 + E[\mathbf{w}^T C^T Q C \mathbf{w}] + E[x_0^T A^T Q A x_0]) \\ &= 2(B^T Q B + R) \mathbf{u} + 2\mu_0 A^T Q B \end{aligned}$$

assuming $(B^T Q B + R)$ is invertible,

$$\mathbf{u} = (B^T Q B + R)^{-1} (\mu_0 A^T Q B)$$

Problem 3

Optimal disposition of a stock. You must sell a total amount, $B > 0$ of a stock in two rounds. In each round you can sell any nonnegative amount of the stock; by the second round all of the initial stock amount B must be sold. The (positive) prices in the two rounds are p_0 and p_1 , respectively. These are independent log-normal variables:

$$\log p_0 \sim \mathcal{N}(\mu_0, \sigma_0^2), \quad \log p_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$$

The goal is to maximize the total expected revenue from the sales in the two rounds. We consider three different information patterns.

- *Prescient.* You know p_0 and p_1 before you decide the amounts to sell in each period.
- *No knowledge.* You do not know the prices, but you know their distribution parameters.
- *Partial knowledge.* You are told the price p_0 before you decide how much to sell in period 0, and you are told the price p_1 before you decide how much to sell in period 1.

a)

PROBLEM:

Find the optimal policies for each of the three different information patterns. The amount sold in each period can depend on the problem data $(B, \mu_0, \mu_1, \sigma_0, \sigma_1)$ and of course the additional information available, which depends on the information pattern.

SOLUTION:

b)

PROBLEM:

Consider the specific case with problem data

$$(B, \mu_0, \mu_1, \sigma_0, \sigma_1) = (100, 0, 0.1, 0.4, 0.4)$$

Plot the distribution of the total revenue for the stochastic control problems for the three different information patterns, using (if necessary) Monte Carlo estimation. Give the expected values of total revenue in each case (again, computed by Monte Carlo estimation).

Hints.

- If $\log x \sim \mathcal{N}(\mu, \sigma^2)$, then $E[x] = e^{\mu + \sigma^2/2}$.
- In matlab you can plot the histogram of a vector v in n bins using the command `hist(v,n)`.

You don't need to know a general method for solving stochastic optimal control problems to solve this problem. You can solve it directly using basic and simple arguments.

SOLUTION:

References

- [1] Jonas R Wagner. Mech 6323 final project.
- [2] Jonas R Wagner. Mech 6327 hw1.