Math 227C: Introduction to Stochastic Differential Equations

Lecturer: Xiaohui Xie
Scribe: Andrew Schaub

Lecture #19

1 Stochastic Control

The focus of this lecture will be on stochastic control. We begin by looking at optimal control in the deterministic case.

1.1 Deterministic Optimal Control

Suppose you have a deterministic ODE

$$\dot{x} = F(x(t), u(t))$$

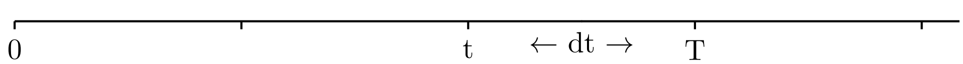
Our goal in using deterministic optimal control is to try to solve the above equation. The variable to be minimized is u over all of it's functions The state variable is $x(t)$, $u(t)$ is the control, and the independent variable is t. We want the system to give rise to optimal performance. So, apply optimal control over a fixed time period $[0, T]$.

$$\min_u \left\{ \int_0^T C[x(t), u(t)]dt + D[X(T)] \right\}$$

The initial state t_0 needs to be close to something, and this will help determine the terminal cost. The goal is to find some optimal use that minimizes the above value function. An issue arrises though upon beginning with $V(x(0), 0)$, and it concerns the identity of u. This issue is a general problem.

1.1.1 Dynamic Programming

The typical way of trying to solve this general problem is using a dynamic programming technique.



At point t there is an optimal value $V(x(t), t)$. One way to try to control the system is by taking small steps and observing how the system evolves. This allows the minimization of the control in some way. Now divide the cost into two components. The first component will be the cost that occurs during the interval dt, and the second component is the cost afterwards.

$$V(x(t), t) = \min_u [C(x(t), u(t))dt] + V(x(t + dt), t + dt)]$$

Essentially the goal is to minimize the above function, and that is the basic idea of dynamic programming. Optimal control is obtained by looking at this small interval, with the difference here focusing on what occurs in the middle. Do a taylor expansion on V, and luckily the taylor expansion will be relatively straightforward. The first order taylor expansion will look like

$$V(x(t + dt), t + dt) = V(x(t), t) + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial X} \cdot \dot{x}(t)dt$$

Taylor expansion will reveal as V terms will cancel eachother out

$$V(x(t+dt),t+dt) = V(x(t),t) + \frac{\partial V}{\partial t} + \frac{\partial V}{\partial X} \cdot \dot{x}(t)dt$$

In order to know that the above taylor series will converge smoothness must be assumed. The third part of the above equation $\frac{\partial V}{\partial X} \cdot \dot{x}(t)dt$ is not necessarily scalar, but will be the dot product between two scalars. The importance of this is because it is recursive, and this will lead to the minimal solution.

$$\min_u \left\{ \frac{\partial V}{\partial t} + \left\langle \frac{\partial V}{\partial X}, F(x(t),u(t)) \right\rangle + C(tx(t),u(t) \right\} = 0$$

The minimum here has to be zero. Move $\frac{\partial V}{\partial t}$ outside,

$$\frac{\partial V}{\partial t} + \min_u \left\{ \left\langle \frac{\partial V}{\partial x}, F(x,u) \right\rangle + C(x,u) \right\} = 0$$

The only cost associated with this is the terminal cost, and that is small. It simply equals

$$V(x,T) = D(x)$$

This is the PDE that needs to be solved to find the optimal control of the system. So the general idea is to try to identify the optimal control first. Once that has been identified and the optimal value has be obtained than the optimal u can be extracted. This needs to be done before solving the complicated PDE, as it has a minimization process subject to the terminal case. This is known as deterministic optimal control, where the best u will be found. So this will lead to the trajectory of x changing, which implies that the cost function will change as well. By doing this though the best control signal will be found to optimize this cost function. The final cost is not independent of initial selections though, as it also depends on D. If a different u is selected, than the boundary conditions of V has to be satisified, and the terminal cost will be satisfied. At the end u is only a function of t. Be warned, this is not like the filtering problem, though this will be more apparent when stochastic control is covered.

These types of equations are known as Hamilton-Jacobian-Bellman (HJB) equations, and are famous if you are studying optimal control. Think about solving this numerically by going backwards. The changes will relate to the difference in the PDE. $\frac{\partial V}{\partial x}$ is known because it is terminal. Many other domains use this technique of dynamic programming by breaking down problems into smaller and smaller portions as well. This is why it is known as deterministic optimal control. The problem is very trivial at t, but as it progresses it solves the original problem. Though unfortunately it is nontrivial to solve.

1.2 Stochastic Optimal Control

To look at a case with stochastic optimal control variance is introduced.

$$dX_t = b(x_t,u_t)dt + \sigma(x_t,u_t)dB_t$$

So where the case in the previous section evolved deterministically, this case evolves as an ODE. Suppose we are assuming a high dimension n, then it belongs to n dimensional.

Therefore it's not really scalar, but is a random variable, by taking domains in the n dimension.

$$X_t \in R^n, \quad B_t : m - \text{dim Brownian motion}$$

$$\sigma(x_t, u_t) \in R^{nxm}$$

So define the value equations and start with $x(t)$ at time t.

$$V(x(t), t) = \min_u \int_t^T C(X(t), u(t))dt + D[X(T)]$$

What is the expectation? Remember $X(T)$ is random,

$$V(x(t), t) = \min_u E\left[\int_t^T C(X(t), u(t))dt + D[X(T)]\right]$$

What type of control is being used? That is the question that needs to be addressed, for there are many types of controls.

1. Open Loop Control (Deterministic Control).

   Suppose $u(t, \omega) = u(t)$. In this case it will be deterministic control (open looped control). Because it's simply fixed as a function of t.

2. Open Looped Control (Feedback Control).

   Suppose, U_t is M_t-adapted, where M_t is the σ-algebra generated by $X_s, 0 \leq S \leq t$. Essentially for this σ-algebra you have all the information about the trajectory from 0 up to a certain time point, the history must be known.

3. Markov Control

   $U(t, \omega) = u_0(t, x_t(\omega))$. Markov control uses less control. There is no history, and there is no memory. It only uses what is current, there is nothing beyond that. An example of where this would be used is a control theory about robots. The robot has to decide to walk or stop, and this decision doesn't depend on the past. Because of this, the Markov control will be the same as the open loop control.

These types of controls explain what type of information is allowed to be used. So regardless of the type of control used in stochastic optimal control we apply the recursive formula, and we have to get expectations.

$$V(x(t), t) + \frac{\partial V}{\partial t} \cdot dt + \frac{\partial V}{\partial x} \cdot dX(t) + \frac{1}{2}(dx_t)^T \frac{\partial^2 V}{\partial x^2}(dx_t)$$

So what is the covariant structure of this $X(t)$. It will simply be the matrix $\sigma(x_t, u_t)$.

$$V(x(t), t) + \frac{\partial V}{\partial t}dt + \left\langle \frac{\partial V}{\partial x}, b(x_t, u_t) \right\rangle dt + \left\langle \frac{\partial V}{\partial x}, \sigma(x_t, u_t)dB_t \right\rangle + \frac{1}{2}\sum_{i,j} a_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j}$$

$$a_{ij} = (\sigma\sigma^T)_{ij}$$

After obtaining the expectation,

$$\frac{\partial V}{\partial t} + \min_u \left\{ < \frac{\partial V}{\partial x}, b(x_t, u_t) > + \frac{1}{2} \sum_{ij} a_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} \right\}$$

Whether or not the solution exists, or is unique are hard questions. In general the goal is a practical solution to the HJB equations. This type of equation does not necessarily have an analytical solution. Though caution needs to be exercised as numerical solutions have issues as well. A question was raised during lecture about the existence of a functional way of approximating the minimum of u using an approximation formula or calculus of variations. There is no clear way of how to do this. Remember, this is a general formula to follow. If the user wants to think about analytical solutions, such as pondering if b is possibly convex, and b is the convex of u. This exploration might lead to a possible solution, but in general it own't be easy.

1.2.1 Example of Linear Stochastic Control

Suppose,

$$dX_t = (H_t X_t + M_t U_t)dt + \sigma_t dB_t$$

In this case,

$$x_0 = x, \quad t \geq 0. \;\; H_t \in R^{n/n}, U_t \in R^k$$

$$\sigma_t R^{n \times m}, \;\; M_t \in R^{n \times k}$$

This minimizes u over the expectation,

$$V^a(x, 0) \min_u E^{x,0} \left\{ \int_0^T (x_t^T C_t X_t + u_t^T D_t u_t) dt + X_T^T R X_T \right\}$$

There are costs associated with control. Try to make the u small. Another cost associated with control is the terminal cost.

$$\psi(t, x) = \min_u V^u$$

This satisfies the partial Jacobian. Add an an s term to make the equation more flexible

$$\frac{\partial \psi}{\partial s} + \min_u \left\{ x^T C_s X + u^T P_s v + \sum_{i=1}^{n} (H_s x + M_s)_i \frac{\partial \psi}{\partial x_i} + \frac{1}{2} \sum_{ij} (\sigma_s \sigma_s^T)_{ij} \frac{\partial^2 \psi}{\partial x_i \partial x_j} \right\} = 0$$

The matrix will determine the quadratic form. We need to specify that matrix. At time $T, S_t = R, a_T = 0$. Solving the second term is the gradient of ψ

$$X^T \dot{S}_t x + \dot{a}_t + \min_u \left\{ x^T C_t x + V^T D_t v + v^T + < H_t x + M_t v, 2S_t x > + \sum_{ij} (\sigma_s \sigma_s^T)_{ij} (S_t)_{ij} = 0 \right\}$$

$$\dot{x} \dot{s}_t x + \dot{a}_t + \min_u \left\{ x^T c_t x + V^T D_t v + < H_t x + M_t v m 2S) t x > \operatorname{tr}[(\sigma_t \sigma_t^T) S_t] \right\} = 0$$

$$2D_t v - 2M_t^T S_t X = 0$$

$$2(S_tX)^T M_t v$$

$$v = -D_t^T M_t^T S_t x$$

$$2(M_t^T S_t x)^T u$$

$$x^T(\dot{s} + C_t - S_t M_t D_E^{-1} M_t^T S_t + 2H_t^T s_t = 0)z + \dot{a}_t + tr(\sigma\sigma^T S)_t = 0$$

1. $\dot{S}_t + C_t - S_t M_t D_t^{-1} M_t^T S_t + 2H_t^T S_t = 0 \;\; S_T = R$
2. $\dot{a}_t = -tr(\sigma\sigma^T s)_t, \;\; a_T = 0$

This type of equation is called a Ricarti equation.

$$\dot{s}_t + S_t A_t S_t + B_t S_t + C_t = 0$$

Find s_t which is an entire matrix. Once you find s_t you scan to find optimal control. This is common in engineering, such as in airplane control. This particular example is an example of linear stochastic control.

2 The Filtering Problem

In principle the key to the filtering problem involves solving an HBJ equation to find the value function V. Linear systems will have a solution.

2.1 General Filtering Problem

The general problem involves a system of equations

$$\text{(System)} \quad dX_t = b(t, x_t)dt + \sigma(t, x_t)dU_t$$

$$\text{(Observations)} \quad dZ_t = C_t, x_t)dt + \gamma(t, x_t)dV_t$$

$$u_t : p\text{-dim brownian motion}$$

$$z_0 = 0V_t : r\text{-dim Brownian motion}$$

Given the observations $\{Z_s\}_{:0\leq s\leq t}$. What's the best estimate $\hat{x}_t$ of the x_t?

1. $\hat{x}_t$ is G_t-measurable, where G_t is the σ-algebra generated by $\{Z_s\}_{0\leq s\leq t}$.
2. By best estimate it's the smallest.

Define K to be

$$K = \{y : \omega- > R^n : Y \text{ is } G_t\text{-measurable}$$

With a finite variance, so $y \in L^2(\omega)$. For those familiar with functional analysis a general functional space.

$$L^2(\omega) = \{X : \omega- > R^n | X \text{ is } L\text{-measurable \& } E[X^2] < \infty$$

Also because it's an L^2 space it defines an inner product as well. This is a Hilbert space, an abstract vector space which possesses an inner product, that allows length and angle measurement.

1. $< x, y >= E[xy]$
2. Norm: $||X||_2 =< x, x >= E[X^2]$

This defines the sample space. K_t is only G_t measurable. When referring to the best solution, what is $\hat{x}_t$. We assume the best solution is

$$E[|x_t - \hat{x}_t|^2] = \inf E[|x_t - y|^2]$$

$$y \in k_t$$

Now comes the important concept. In Hilbert space these are close to subspace. x_t is a random variable, itself does not belong to k_t, so the geometric meaning is that it is the projection, and not just a projection, but the orthogonal projection. $\hat{x}_t$ is the orthogonal projection of x onto k_t. This leads to the minimal distance from a closed subspace. To reiterate the importance of this in geometric terms. The best estimate is the filtering of the orthogonal projection of x_t onto k_t all functions are g_t measurable. L is normal defined. These conceptual changes are very important to grasp.

... To be covered in the next and final lecture