1. EXAMPLES OF DETERMINISTIC CONTROLLED MODELS. THE BELLMAN EQUATION

The course of "Stochastic Control" includes the following topics:

1. Examples of controlled deterministic models: Linear quadratic Deterministic problem for finite and finite horizons. The Bellman equation;

2. Stochastic Calculus: Wiener process, Stochastic Itô integral, Itô formula, Doob inequality, Stochastic Itô equation, Fokker-Plank- Kolmogorov (FPK) equation;

3. Filtering: Kalman filter for continuous and discrete time cases, Generalized Kalman filter, Conditionally Gaussian filter, Extended Kalman filter;

4. Examples of transmission Gaussian signal through channel with noiseless feedback;

4. Stochastic feedback control (complete and incomplete observations), Control of noise intensity, Stationary regimes;

5. Examples of stochastic control in finance.

1. **Deterministic dynamic controlled model.**

Example of open loop control. We consider a simplest dynamic model described by a linear differential equation

$$\dot{X}_t = aX_t + u(t), \tag{1.1}$$

subject to fixed initial point $X_0 = x$.

Here the function $u(t)$ is named "control action" or "control" and X_t "controlled process". Applying control $u(t)$ during some time interval we may move X_t from the initial point to some new positions. For example, we apply control $u(t)$, $0 \le t \le T$ and need to move X_t to from $X_0 = x$ to a new position

$$X_T = y. \tag{1.2}$$

A control guaranteeing that move is not unique and from practical point of view it makes sense to use some "economical" version of control. To realize this comment, let us introduce the cost function

$$J(u) := \int_0^T u^2(t)dt. \tag{1.3}$$

We intend to find such a control $u^\circ(t)$ which guarantees (1.2) and for which the cost function $J(u^\circ)$ is minimal. In other words,

$$J(u^\circ) \le J(u)$$

for any control $u(t)$ providing (1.2).

There are a few types of controls:

- $\boxed{\text{open loop control}}$ when $u(t)$ is a function of the time parameter t only;
- $\boxed{\text{feedback control}}$ when $u(t) = u(t, X_t)$;
- $\boxed{\text{past-dependent feedback control}}$ when $u(t) = u(t, X_{[0,t]})$, here $X_{[0,t]} = \{X_s, 0 \leq s \leq t\}$.

For our example we choose open loop control and find the optimal control $u^\circ(t)$.

Step 1: Cost functional lower bound. We use that the equation given in (1.2) is linear and find its solution

$$X_T = e^{aT}\Big(x + \int_0^T e^{-at}u(t)dt\Big)$$

for some arbitrary control $u(t)$. Among all controls we choose admissible controls, that is such that $X_T = y$. In other words, if $u(t)$ is admissible control, then we have $y = e^{aT}\Big(x + \int_0^T e^{-at}u(t)dt\Big)$ and moreover

$$\int_0^T e^{-at}u(t)dt = \frac{y - e^{aT}x}{e^{aT}}. \tag{1.4}$$

Taking into account the form of $J(u)$, we apply the Cauchy-Schwarz inequality

$$\Big(\int_0^T e^{-at}u(t)dt\Big)^2 \leq \int_0^T u^2(t)dt \int_0^T e^{-2at}dt := J(u) \int_0^T e^{-2at}dt$$

and find the following lower bound for the cost function:

$$J(u) \geq \frac{\big(y - e^{aT}x\big)^2}{e^{2aT}\big(\int_0^T e^{-2as}ds\big)}. \tag{1.5}$$

Step 2: Creating the optimal control. Obviously, if there exists an admissible control $\widetilde{u}(t)$ for which the lower bound given in (1.5) is attainable, i.e.

$$J(\widetilde{u}) = \frac{\big(y - e^{aT}x\big)^2}{e^{2aT}\big(\int_0^T e^{-2as}ds\big)} \quad \Bigg(= 2a\frac{\big(y - e^{aT}x\big)^2}{e^{2aT} - 1}\Bigg).$$

then $\widetilde{u}(t)$ is the optimal control: $\widetilde{u}(t) \equiv u^\circ(t)$.

Notice that for the inequality in (1.5) to be equality it suffices $\widetilde{u}(t)$ is in a proportion to e^{-at}:

$$\widetilde{u}(t) = Ce^{-at},$$

and

$$\Big(\int_0^T e^{-at}\widetilde{u}(t)dt\Big)^2 = \int_0^T (\widetilde{u}(t))^2dt \int_0^T e^{-2at}dt$$

that is

$$\left(\int_0^T e^{-at}\widetilde{u}(t)dt\right)^2 = C^2\left(\int_0^T e^{-2at}dt\right)^2.$$

On the other hand, since

$$\int_0^T e^{-at}\widetilde{u}(t)dt = \frac{y - e^{aT}x}{e^{aT}},$$

we find

$$C = \frac{y - e^{aT}x}{e^{aT}\int_0^T e^{-2as}ds}.$$

Thus

$$\widetilde{u}(t) = \frac{y - e^{aT}x}{e^{aT}\int_0^T e^{-2as}ds}e^{-at} = 2a\frac{y - e^{aT}x}{e^{aT}(1 - e^{-2aT})}e^{-at}. \tag{1.6}$$

Example of closed loop (feedback) control.

Assume X_t is defined again by (1.1) and now

$$J(x,u) = X_T^2 + \int_0^T \big(X_t^2 + u^2(t)\big)dt \tag{1.7}$$

is the cost function. The main difference of (1.7) from (1.3) consists in what we need "to pay" not only for $u^2(t)$ but also for $X^2(t)$. In other words, we have, by choosing an appropriate control action, to minimize not only $\int_0^T u^2(t)dt$ but also $X_T^2 + \int_0^T X_t^2dt$. Both tasks contradict to each other.

Nevertheless the *feedback* optimal control exists and we find it.

1. Assume $x = 0$. Then, obviously, $u(t) \equiv 0$ is the optimal control, $X_t \equiv 0$ and $J(0,0) = 0$.

2. If $x \neq 0$, to find the optimal feedback control we apply *Dynamical Programming Principle* (DPP) introduced by R. Bellman. Set

$$J(t,x,u) = X_T^2 + \int_t^T \big(X_s^2 + u^2(s)\big)ds, \ 0 \le t < T, \tag{1.8}$$

where X_s solves the differential equation

$$\dot{X}_s = aX_s + u(s), \ t < s \le T, \tag{1.9}$$

subject to $X_t = x$, and

$$V(t,x) = \inf_{u(s), t\le s\le T} J(t,u,x)$$

named the Bellman function. Assume for a moment that the optimal control exists $\widetilde{u}(s)$, $t \le s \le T$ exists for any $t < T$, that is the Bellman function is well defined.

For the example considered, the dynamical programming principle postulates: for every $0 \le t < t' \le T$

$$
\begin{aligned}
V(t,x) &= \inf_{u(s),0\le s\le T}\Big\{X_T^2+\int_0^{t'}\big(X_s^2+u^2(s)\big)ds+\int_{t'}^{T}\big(X_s^2+u^2(s)\big)ds\Big\} \\
&= \inf_{u(s'),t\le s'\le t'}\Big\{\int_t^{t'}(X_{s'}^2+u^2(s'))ds+\inf_{u(s),t'\le s\le T}\Big[X_T^2+\int_{t'}^{T}\big(X_s^2+u^2(s)\big)ds\Big]\Big\} \\
&= \inf_{u(s'),t\le s'\le t'}\Big\{\int_t^{t'}(X_{s'}^2+u^2(s'))ds+V(t',X_{t'})\Big\}.
\end{aligned}
\tag{1.10}
$$

In words: if the optimal control $u^*(s)$, $t' \le s \le T$ is known for any initial point $X_{t'}, t'$ and the Bellman function $V(t', X_{t'})$ is known, then the optimal control $u^*(s')$, $t \le s' \le t'$ is defined from (1.10), where

$$\dot{X}_{s'} = aX_{s'} + u(s')$$

and $X_{t'}$, involved in $V(t', X_{t'})$, is defined from the above differential equation.

DPP provides *the Bellman equation.* To derive it, set $\delta = t' - t$. Taking small δ and assuming $V(t, x)$ is smooth in t, x write

$$
\begin{aligned}
V(t,x) &= \inf_u\big\{(x^2+u^2)\delta+V(t+\delta,X_{t+\delta}\big\}+o(\delta) \\
&= \inf_u\big\{(x^2+u^2)\delta+V(t,x)+V_t(t,x)\delta+V_x(t,x)(ax+u)\delta\big\}+o(\delta).
\end{aligned}
$$

Then, passing $\delta \to 0$, we get

$$-\frac{\partial V(t,x)}{\partial t}=\inf_u\Big\{x^2+u^2+\frac{\partial V(t,x)}{\partial x}\big[ax+u\big]\Big\} \tag{1.11}$$

the Bellman equation subject to boundary condition $V(T, x) = x^2$.

How to use the Bellman equation? Notice first that

$$V(0,x)=\inf_u J(x,u),$$

that is $V(0, x)$ is the optimal value of the cost function. Now, we discuss a role of the function

$$u^*(t,x)=\operatorname{argmin}\Big\{x^2+u^2+\frac{\partial V(t,x)}{\partial x}\big[ax+u\big]\Big\}. \tag{1.12}$$

Proposition 1. *$u^*(t, X_t^*)$, with X_t^* defined by a differential equation*

$$\dot{X}_t^* = aX_t^* + u^*(t, X_t^*)$$

subject to $X_0^* = x$, is the optimal control. Moreover $u^*(t, x) = -\Gamma(t)x^2$.

Proof. A minimization in u of the quadratic form $x^2+u^2+\frac{\partial V(t,x)}{\partial x}\big[ax+u\big]$ provides

$$u^*(t,x)=-\frac{1}{2}\frac{\partial V(t,x)}{\partial x}; \tag{1.13}$$

notice that

$$0=\frac{\partial}{\partial}\Big(x^2+u^2+\frac{\partial V(t,x)}{\partial x}\big[ax+u\big]\Big)=2u+\frac{\partial V(t,x)}{\partial x}.$$

So, $u^*(t,x)$ (1.11) is transformed to

$$-\frac{\partial V(t,x)}{\partial t}=\Big\{x^2+\frac{\partial V(t,x)}{\partial x}ax-\frac{1}{4}\Big(\frac{\partial V(t,x)}{\partial x}\Big)^2\Big\} \tag{1.14}$$

subject to the boundary condition $V(T,x)=x^2$.

Notice also that a quadratic form in x might be considered as a candidate to be a solution of (1.14). So, we taste the function

$$\begin{aligned}\widetilde{V}(t,x)&=\Gamma(t)x^2+B(t)x+Q(t)\\ \widetilde{V}(T,x)&=V(T,x)\end{aligned} \tag{1.15}$$

whether it is a solution of (1.14). In addition, we assume $\Gamma(t),B(t),Q(t)$ are differentiable functions. Substituting $\widetilde{V}(t,x)$ in (1.14), for any t we get the identity in x:

$$-\Big(\dot{\Gamma}(t)x^2+\dot{B}(t)x+\dot{Q}(t)\Big)\equiv x^2+ax\big(2\Gamma(t)x+B(t)\big)-\frac{1}{4}\big(2\Gamma(t)x+B(t)\big)^2$$

which provides the differential equations for $\Gamma(t)$, $B(t)$ and $Q(t)$:

$$\begin{aligned}-\dot{\Gamma}(t)&=1+2a\Gamma(t)-\Gamma^2(t)\\ -\dot{B}(t)&=aB(t)-2\Gamma(t)B(t)\\ -\dot{Q}(t)&=-\frac{1}{4}B^2(t).\end{aligned} \tag{1.16}$$

Moreover, the second part in (1.15) gives

$$\Gamma(T)x^2+B(T)x+Q(T)=x^2.$$

Hence the boundary conditions for (1.16) are the following:

$$\Gamma(T)=1,\ \ B(T)=0,\ \ Q(T)=0. \tag{1.17}$$

Let us solve the differential equation (linear) for $B(t)$. Write

$$B(t)=B(0)\exp\Big(\int_0^t(a-2\Gamma(s))ds\Big).$$

Since $B(T)=0$, we have $B(0)=0$ and thus $B(t)\equiv 0$. We can conclude now that also $Q(t)\equiv 0$.

Hence, we have

$$\widetilde{V}(t,x) = \Gamma(t)x^2 \tag{1.18}$$

is one of solution of the Bellman equation.

Let us take $u(t)$ some admissible control and let X_t be a corresponding controlled process: $\dot{X}_t = aX_t + u(t)$ and $J(x,u) < \infty$.

From the definition of $u^*(t,x)$ it follows that for any (t,x)

$$x^2 + u^2(t) + \frac{\partial \widetilde{V}(t,x)}{\partial x}\big[ax + u(t)\big] \le \frac{\partial \widetilde{V}(t,x)}{\partial t}.$$

Using the above inequality and computing $\frac{d\widetilde{V}(t,X_t)}{dt}$ we find

$$\begin{aligned}
\frac{d\widetilde{V}(t,X_t)}{dt} &= \frac{\partial \widetilde{V}(t,X_t)}{\partial t} + \frac{\partial \widetilde{V}(t,X_t)}{\partial x}[aX_t + u(t)] \\
&\ge -\frac{\partial \widetilde{V}(t,X_t)}{\partial x}[aX_t + u(t)] - [X_t^2 + u^2(t)] + \frac{\partial \widetilde{V}(t,X_t)}{\partial x}[aX_t + u(t)] \\
&= -[X_t^2 + u^2(t)]. \qquad (1.19)
\end{aligned}$$

Hence, taking into account that $\widetilde{V}(T,x) = x^2$, we find

$$X_t^2 = \widetilde{V}(T,X_T) \ge \widetilde{V}(0,x) - \int_0^T \Big(X_t^2 + u^2(t)\Big)dt.$$

Consequently, $\widetilde{V}(0,x) \le X_T^2 + \int_0^T \Big(X_t^2 + u^2(t)\Big)dt := J(x,u)$ i.e.

$$J(x,u) \ge \widetilde{V}(0,x) = \Gamma(0)x^2. \tag{1.20}$$

Now, taking $u(t) \equiv u^*(t,X_t^*)$ and taking into account that then (1.19) is replaced by

$$\begin{aligned}
\frac{d\widetilde{V}(t,X_t^*)}{dt} &= \frac{\partial \widetilde{V}(t,X_t^*)}{\partial t} + \frac{\partial \widetilde{V}(t,X_t^*)}{\partial x}[aX_t^* + u^*(t,X_t^*)] \\
&= -\frac{\partial \widetilde{V}(t,X_t^*)}{\partial x}[aX_t^* + u^*(t,X_t^*)] - \big[(X_t^*)^2 + (u^*(t,X_t^*))^2\big] \\
&\quad + \frac{\partial \widetilde{V}(t,X_t^*)}{\partial x}[aX_t^* + u^*(t,X_t^*)] \\
&= -\big[(X_t^*)^2 + (u^*(t,X_t^*)^2(t)\big], \qquad (1.21)
\end{aligned}$$

we get

$$J(x,u^*) = \widetilde{V}(0,x) = \Gamma(0)x^2. \tag{1.22}$$

Now, comparing (1.20) and (1.22) we claim that $u^*(t,X_t^*)$ is the optimal control.

To define the function $u^*(t,x)$ explicitly, we use formula (1.13) with V replaced by $\widetilde{V}$ and formula (1.18). Then $u^*(t,x) = -\Gamma(t)x^2$. □

2. **Home work:** **Discrete time case. Feedback control. Bellman equation**

Let $X_k, k = 0, 1, ..., N$ be a controlled sequence defined by linear recursion

$$X_{k+1} = aX_k + cu_k$$

subject to fixed initial condition $X_0 = x$, where $u_k, k = 0, 1..., N-1$ is control actions.

Find the optimal control minimizing the cost functional ($p \geq 0$, $q > 0$):

$$J(x, u) = pX_N^2 + \sum_{k=1}^{N-1} (pX_k^2 + qu_k^2). \tag{2.1}$$

1. Derive the Bellman equation for the Bellman function

$$V(x, n) = \min_{u_k, n \leq N} \Big[pX_N^2 + \sum_{k=1}^{N-1} (pX_k^2 + qu_k^2) \Big],$$

where $X_k, n \leq k \leq n$ is defined by recursion (2.1) with $X_n = x$.

2. Find the optimal control.