10. STOCHASTIC FEEDBACK CONTROL.

10.1. **Controlled drift and diffusion.**

In the previous lecture, the control affects a drift of controlled process. Here, we consider a controlled model with a control affecting drift and intensity of a white noise.

We will analyze the following model:

$$dX_t = [a_0(t)X_t dt + b_0(t)dW_t] + u(t)[a_1(t)dt + b_1(t)dW'_t] \tag{10.1}$$

subject to the fixed initial point $X_0 = x$, where W_t and W'_t are independent Wiener processes and $a_0(t), a_1(t), b_0(t), b_1(t)$ are known (bounded) functions. The control $u(t)$ is taken in feedback form: $u(t) = u(t, X_{[0,t]})$ and has to be chosen such that to minimize the quadratic cost functional

$$J(x,u) = \boldsymbol{E}\Big\{rX_T^2 + \int_0^T \Big(p(t)X_t^2 + q(t)u^2(t, X_{[0,t]})\Big)dt\Big\}, \tag{10.2}$$

where $r \geq 0$, $p(t)$ and $q(t)$ are continuous function, $p(t) \geq 0$ and

$$q(t) \geq c > 0.$$

10.2. **Bellman equation.**

Introduce the Bellman function

$$V(t,x) = \inf_{u(s), t \leq s \leq T} \boldsymbol{E}\Big\{rX_T^2 + \int_0^T \Big(p(s)X_s^2 + q(t)u^2(s, X_t^s)\Big)ds\Big\},$$

where X_s is defined for $s > t$ and $X_t = x$

$$dX_s = [a_0(s)X_s ds + b_0(s)dW_s] + u(s)[a_1(s)ds + b_1(s)dW'_s], \ s > t$$

and the Bellman equation:

$$\begin{aligned} -V'_t(t,x) \ &= \ \inf_u \Big(\frac{1}{2}[b_0^2 + b_1^2(t)u^2]V''_{xx}(t,x) + \\ &\qquad +[a_0(t)x + a_1(t)u]V'_x(t,x) + p(t)x^2 + q(t)u^2\Big) \end{aligned} \tag{10.3}$$

subject with boundary condition $V(T,x) = rx^2$. The parameter u, minimizing the quadratic form

$$Q(u) = \frac{1}{2}b_1^2(t)u^2V''_{xx}(t,x) + a_1(t)uV'_x(t,x) + q(t)u^2,$$

is defined by the linear equation

$$[b_1^2(t)V''_{xx}(t,x) + 2q(t)]u + a_1(t)V'_x(t,x) = 0.$$

Latter, it will be shown that

$$V''_{xx}(t,x) \geq 0. \tag{10.4}$$

Under (10.4) we have $[b_1^2(t)V''_{xx}(t,x) + 2q(t)] > 0$ and so the minimum of $Q(u)$ is attainable at the point

$$u^\circ(t,x) = -\frac{a_1(t)V'_x(t,x)}{b_1^2(t)V''_{xx}(t,x) + 2q(t)}.$$

The substitution of $u^\circ(t,x)$ in the Bellman equation imply the second order non linear partial differential equation:

$$\begin{aligned} -V'_t(t,x) &= \Big[\frac{1}{2}b_0^2(t)V''_{xx}(t,x) + a_0(t)xV'_x(t,x) + p(t)x^2\Big] \\ &\quad -\frac{1}{2}\frac{\Big(a_1(t)V'_x(t,x)\Big)^2}{b_1^2(t)V''_{xx}(t,x) + 2q(t)}\Big] \end{aligned} \tag{10.5}$$

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

As for the LQG model, we will find a solution of the Bellman equation as a quadratic in x function:

$$V(t,x) = \Gamma(t)x^2 + B(t)x + Q(t)$$

with coefficients $\Gamma(t), B(t), Q(t)$ continuously differentiable in t. The boundary condition gives

$$\Gamma(T) = r, \quad B(T) = 0, \quad Q(T) = 0.$$

Further, differentiating $V(t,x) := \Gamma(t)x^2 + B(t)x + Q(t)$ in t we obtain

$$\begin{aligned} \frac{\partial V(t,x)}{\partial x} &= 2\Gamma(t)x + B(t) \\ \frac{\partial^2 V(t,x)}{\partial x^2} &= 2\Gamma(t). \end{aligned}$$

With these derivatives, (10.5) is transformed to

$$\begin{aligned} -\Big(\dot\Gamma(t)x^2 + \dot B(t)x + \dot Q(t)\Big) &= \Big[b_0^2(t)\Gamma(t) + 2a_0(t)x(\Gamma(t)x + B(t)) + p(t)x^2\Big] \\ &\quad -\frac{1}{2}\frac{\Big(2a_1(t)(\Gamma(t)x + B(t))\Big)^2}{2b_1^2(t)\Gamma(t) + 2q(t)}\Big] \end{aligned}$$

and, due to an arbitrariness of x, we obtain the ordinary differential equations for $\Gamma(t)$, $B(t)$, and $Q(t)$:

$$\begin{aligned} -\dot\Gamma(t) &= 2a_0(t)\Gamma(t) + p(t) - \frac{2a_1^2(t)\Gamma^2}{2b_1^2(t)\Gamma(t) + 2q(t)} \\ -\dot B(t) &= 2a_0(t)B(t) - 4\frac{a_1^2(t)\Gamma(t)}{2b_1^2(t)\Gamma(t) + 2q(t)}B(t) \\ -\dot Q(t) &= b_0^2(t)\Gamma(t) - 2\frac{a_1^2(t)B^2(t)}{2b_1^2(t)\Gamma(t) + 2q(t)}. \end{aligned}$$

The second equation is linear and subject to "zero" boundary condition. Therefore $B(t) \equiv 0$. In this case

$$Q(t) = \int_t^T b_0^2(s)\Gamma(s)ds.$$

Thus

$$V(t,x) = \Gamma(t)x^2 + Q(t).$$

Notice that in the standard LQG problem $\Gamma(t)$ solves the Riccati equation. In the case considered the differential equation for $\Gamma(t)$ is not the Riccati equation (it becomes the Riccati equation only under $b_1(t) \equiv 0$).

Let us show that (10.4) holds true. Since

$$V''_{xx}(t,x) \equiv 2\Gamma(t)$$

we have show that $\Gamma(t) \geq 0$. Set $\overline{\Gamma}(t) = \Gamma(T-t)$ and note that on the time interval $[0,T]$, the function $\overline{\Gamma}(t)$ is the solution of the forward differential equation (in contrast to $\Gamma(t)$ being the solution of the backward one):

$$\dot{\overline{\Gamma}}(t) = 2a_0(T-t)\overline{\Gamma}(t) + p(T-t) - \frac{2a_1^2(T-t)\overline{\Gamma}^2(t)}{2b_1^2(T-t)\overline{\Gamma}(t) + 2q(T-t)}$$

where $\overline{\Gamma}(0) = r \geq 0(!)$. Assume for a moment that $\overline{\Gamma}(t)$ might be negative at some time value τ: $\overline{\Gamma}(\tau) = 0$. Since $p(T-\tau) \geq 0$, we find that $\dot{\overline{\Gamma}}(\tau) = 0$. Hence $\overline{\Gamma}(t)$ is "reflected at zero" and so remains nonnegative.

10.3. **The optimal control.**

Now, we are in the position to define $u^\circ = u^\circ(t,x)$:

$$u^\circ(t,x) = -\frac{2a_1(t)\Gamma(t)}{2b_1^2(t)\Gamma(t) + 2q(t)}x.$$

The function $u^\circ(t,x)$ is linear in x with bounded gain

$$\max_{0\leq t\leq T} \frac{2|a_1(t)|\Gamma(t)}{2b_1^2(t)\Gamma(t) + 2q(t)} < \infty$$

so that the differential equation

$$dX_t^\circ = [a_0(t)X_t^\circ dt + b_0(t)dW_t] + u^\circ(t,X_t^\circ)[a_1(t)dt + b_1(t)dW'_t]$$

obeys the unique solution and $u^\circ(t,X_t^\circ)$ is an admissible Markov control.

Theorem. *The control $u^\circ(t,X_t^\circ)$ is optimal.*

Proof: Let us show that $J(x,u^\circ) = V(0,x)$. By the Itô formula applied to $V(t,X_t^\circ)$ we find

$$
\begin{aligned}
V(T,X_T^\circ) &= V(0,x)+\int_0^T[V_t'(t,X_t^\circ)dt+V_x'(t,X_t^\circ)dX_t^\circ]dt \\
&+\frac{1}{2}\int_0^T[b_0^2(t)+b_1^2(t)]V_{xx}''(t,X_t^\circ)dt.
\end{aligned}
$$

On the other side, due to the Bellman equation

$$
\begin{aligned}
V_t'(t,X_t^\circ) &= -\Big(\frac{1}{2}[b_0^2+b_1^2(t)(u^\circ(t,X_t^\circ))^2](t)V_{xx}''(t,X_t^\circ)+ \\
&+[a_0(t)X_t^\circ+a_1(t)u]V_x'(t,X_t^\circ)+p(t)(X_t^\circ)^2+q(t)(u^\circ(t,X_t^\circ))^2\Big).
\end{aligned}
$$

Further, by virtue of the Itô equation for X_t° we have

$V_x'(t,X_t^\circ)dX_t^\circ = V_x'(t,X_t^\circ)[a_0(t)X_t^\circ dt+b_0(t)dW_t]+u^\circ(t,X_t^\circ)[a_1(t)dt+b_1(t)dW_t'].$

Therefore, we obtain

$$
\begin{aligned}
V_t'(t,X_t^\circ)dt+V_x'(t,X_t^\circ)dX_t^\circ &= -[p(t)(X_t^\circ)^2+q(t)(u^\circ(t,X_t^\circ))^2]dt \\
&+b_0(t)dW_t+b_1(t)dW_t'.
\end{aligned}
$$

Finally, by the boundary condition $V(T,X_T^\circ)=rX_T^\circ$, we find

$$
rX_T^\circ+\int_0^T[p(t)(X_t^\circ)^2+q(t)(u^\circ(t,X_t^\circ))^2]dt = V(0,x)+\int_0^T[b_0(t)dW_t+b_1(t)dW_t']
$$

and, taking the expectation from both parts, we get

$$
J(x,u^\circ)=V(0,x).
$$

Let us show that for any admissible control $\widetilde{u}(t)$

$$
J(x,\widetilde{u})\geq V(0,x) \tag{10.6}
$$

Denote by $\widetilde{X}(t)$ the controlled process generated by this control:

$$
d\widetilde{X}_t=[a_0(t)\widetilde{X}_t dt+b_0(t)dW_t]+\widetilde{u}(t)[a_1(t)dt+b_1(t)dW_t']. \tag{10.7}
$$

By (10.3), we have

$$
\begin{aligned}
-V_t'(t,\widetilde{X}_t) &\leq \Big(\frac{1}{2}[b_0^2+b_1^2(t)(\widetilde{u}(t))^2](t)V_{xx}''(t,\widetilde{X}_t)+ \\
&+ [a_0(t)\widetilde{X}_t+a_1(t)\widetilde{u}(t)]V_x'(t,\widetilde{X}_t)+p(t)(\widetilde{X}_t)^2+q(t)(\widetilde{u}(t))^2\Big)
\end{aligned} \tag{10.8}
$$

The Itô formula, applied to $V(t,\widetilde{X}_t)$, and the equality $V(T,\widetilde{X}_T)=r\widetilde{X}_T^2$ provide

$$
\begin{aligned}
r\widetilde{X}_T^2 &= V(0,x) + \int_0^t \Big(V_t'(t,\widetilde{X}_t)dt + V_x'(t,\widetilde{X}_t)[b_0(t)dW_t + \widetilde{u}(t)b_1(t)dW_t'] \\
&\quad + \frac{1}{2}V_{xx}''(t,\widetilde{X}_t)[b_0^2(t) + (\widetilde{u}(t))^2 b_1^2(t)]dt \Big) \\
&\geq V(0,x) - \int_0^T [p(t)(\widetilde{X}_t)^2 + q(t)(\widetilde{u}(t))^2]dt \\
&\quad + \int_0^T V_x'(t,\widetilde{X}_t)[b_0(t)dW_t + \widetilde{u}(t)b_1(t)dW_t'],
\end{aligned}
$$

that is

$$
\begin{aligned}
& r\widetilde{X}_T^2 + \int_0^T [p(t)(\widetilde{X}_t)^2 + q(t)(\widetilde{u}(t))^2]dt \\
\geq\ & V(0,x) + \int_0^T V_x'(t,\widetilde{X}_t)[b_0(t)dW_t + \widetilde{u}(t)b_1(t)dW_t'].
\end{aligned}
$$

Consequently

$$
J(x,\widetilde{u}) := \boldsymbol{E}\Big(r\widetilde{X}_T^2 + \int_0^T [p(t)(\widetilde{X}_t)^2 + q(t)(\widetilde{u}(t))^2]dt \Big) \geq V(0,x).
$$

The theorem assertion follows now from an arbitrariness of $\widetilde{u}$. □

10.4. **Remarks and example.**

1. The differential equation for $\Gamma(t)$ is not Riccati equation.
2. The controlled process corresponding to the optimal control is not obligatory Gaussian .

Let us consider the model, $X_0 = x$ and

$$
dX_t = u(t)[dt + dW_t']. \tag{10.9}
$$

The optimal control

$$
u^\circ(t, X_t^\circ) = -\frac{2\Gamma(t)}{2\Gamma(t) + 2q(t)} X_t^\circ
$$

and therefore, with $\alpha(t) = -\frac{2\Gamma(t)}{2\Gamma(t)+2q(t)}$,

$$
dX_t^\circ = \alpha(t) X_t^\circ [dt + dW_t'].
$$

Applying the Itô formula we find

$$
X_t^\circ = x \exp\Big(\int_0^t [\alpha(s) - \frac{1}{2}\alpha^2(s)]ds + \int_0^t \alpha(s) dW_s'. \Big)
$$

Hence X_t° is a logarithmic function of Gaussian process and so the X_t° is not Gaussian process.

Thus, this model is not LGQ (we lost G-(Gaussiness) and have only LQ-model (L-linear, Q-quadratic).

Exercise. Let $\psi(t)$ is a smooth trajectory which has being tracked by a controlled process $X_0 = x$,

$$dX_t = [aX_t + cu(t)]dt + bdW_t.$$

This requirement is reflected in the cost functional

$$J(x,u) = \boldsymbol{E}\Big\{(X_T - \psi(T))^2 + \int_0^T \big\{(X_t - \psi(t))^2 + u^2(t)\big\}dt\Big\}.$$

To apply for this case a general method of creating the optimal control, we introduce a new controlled process $\widetilde{X}_t = X_t - \psi(t)$ and note that

$$J(x,u) = \boldsymbol{E}\Big\{\widetilde{X}_t^2 + \int_0^T \big\{\widetilde{X}_t^2 + u^2(t)\big\}dt\Big\}.$$

At all, $\widetilde{X}_0 = x - \psi(0)$ and

$$\widetilde{X}_t = [-\dot{\psi}(t) + a\psi(t) + a\widetilde{X}_t]dt + bdW_t.$$

The optimal control for new model is known and defined as $\widetilde{u}^\circ(t) = -c\Gamma(t)\widetilde{X}_t^\circ$, where $\Gamma(t)$ solves corresponding Riccati's equation. Hence the optimal control for the original problem is: $u^\circ(t) = -c\Gamma(t)(X_t^\circ - \psi(t))$ with

$$dX_t^\circ = [aX_t^\circ - c^2\Gamma(t)(X_t^\circ - \psi(t))]dt + bdW_t.$$

Home work.

1. Find the optimal control for the controlled process

$$dX_t = [aX_t + u(t)]dt + bX_t dW_t$$

and the cost functional

$$J(x,u) = \boldsymbol{E}\Big\{rX_T^2 + \int_0^T [X_t^2 + u^2(t)]dt\Big\}.$$

Hint: Use the Bellman equation:

$$-V_t' = \min_u \Big(V_x'[ax+u] + \frac{1}{2}b^2x^2V_{xx}'' + x^2 + u^2\Big)$$

subject to $V(T,x) = x^2$. Find the Bellman function as a quadratic form in x.