Dynamic Programming for General Linear Quadratic Optimal Stochastic Control with Random Coefficients

Shanjian Tang

Fudan University, Shanghai

Beijing Institute of Technology, Beijing
July 2, 2015

Outline

1. Formulation of stochastic linear quadratic control problems

2. The case of deterministic coefficients

3. The case of random coefficients

4. The method of stochastic maximum principle

5. Dynamic programming approach to backward stochastic Riccati equation

1. Formulation of linear quadratic optimal stochastic control problem with random coefficients

The controlled stochastic system:

$$dX_t = (A_t X_t + B_t u_t)\, dt + (C_t X_t + D_t u_t)\, dW_t, \quad X_s = x$$

The cost functional

$$J(u; s, x) = \frac{1}{2} E\left\{ \langle M X_T, X_T \rangle + \int_s^T [\langle Q_t X_t, X_t \rangle + \langle N_t u_t, u_t \rangle]\, dt \,\middle|\, \mathscr{F}_s \right\}$$

Assume that

$$A, B, C, D, Q, N :\in \mathscr{L}^\infty_{\mathscr{F}}, \quad M \in L^\infty(\Omega, \mathscr{F}_T, P; R^{n\times n})$$

and

$$Q \geq 0, \quad M \geq 0, \quad N \geq \delta I_{m\times m}.$$

The optimal control problem: find $\bar{u} \in \mathscr{L}^2_{\mathscr{F}}(s, T; R^m)$ such that

$$J(\bar{u}; s, x) = \operatorname*{ess.inf}_{u\in \mathscr{L}^2_{\mathscr{F}}(s,T;R^m)} J(u; s, x) =: V(s, x).$$

2. The case of deterministic coefficients

Wonham (1968, sicon)

$$\begin{cases} -K_t' = A_t^* K_t + K_t A_t + C_t^* K_t C_t + Q_t - G(t, K_t), \\ K_T = M. \end{cases}$$

The function G is nonlinear:

$$G(t, K) = (KB_t + C_t^* K D_t)(N_t + D_t^* K D_t)^{-1}(KB_t + C_t^* K D_t)^*.$$

The optimal control

$$\overline{u}_t = -(N_t + D_t^* K D_t)^{-1}(KB_t + C_t^* K D_t)^* \overline{X}_t.$$

The value function $V(s, x) = \langle K_s x, x \rangle, \quad (s, x) \in [0, T] \times \mathbb{R}^n.$

Solution of Riccati equation: three methods

(i) direct methods (Picard iteration, Newton's successive approximation)

(ii) variational method

(iii) dynamic programming method

(i) Bellman's quasi-linearization: Newton's algorithm

The optimal feedback control law

$$\bar{u}(t) = -(N_t + D_t^* K_t D_t)^{-1}(K_t B_t + C_t^* K_t D_t)^* \bar{X}_t.$$

Define

$$F(t, K) := -(N_t + D_t^* K D_t)^{-1}(K B_t + C_t^* K D_t)^*,$$

$$\widehat{Q}_t(K) := Q_t + F^*(t, K) N_t F(t, K),$$

$$\widehat{A}_t(K) := A_t + B_t F(t, K), \quad \widehat{C}_t(K) := C_t + D_t F(t, K).$$

$$-K_t' = \widehat{A}_t^*(K_t) K_t + K_t \widehat{A}_t(K_t) + \widehat{C}_t^*(K_t) K_t \widehat{C}_t(K_t) + \widehat{Q}_t(K_t), \quad K_T = M.$$

Iterating algorithm (due to Wonham (sicon, 1968)): Set $K_t^0 \equiv M$, and for $i = 1, 2, \ldots$,

$$\begin{aligned} -\frac{dK_t^i}{dt} &= \widehat{A}_t^*(K_t^{i-1}) K_t^i + K_t^i \widehat{A}_t(K_t^{i-1}) + \widehat{C}_t^*(K_t^{i-1}) K_t^i \widehat{C}_t(K_t^{i-1}) + \widehat{Q}_t(K_t^{i-1}) \\ &= A_t^* K_t^i + K_t^i A_t + C_t^* K_t^i C_t + Q_t - G(t, K_t^{i-1}) - G_K(t, K_t^{i-1})(K_t^i - K_t^{i-1}), \\ K_T^i &= M. \end{aligned}$$

Then, we have

$$K_t^i \geq 0, \quad K_t^i \geq K_t^{i+1}, \quad K_t := \lim_{i \to \infty} K_t^i.$$

(ii) The method of stochastic maximum principle (SMP)

$$\begin{cases} dX_t = (A_tX_t + B_tu_t)\,dt + (C_tX_t + D_tu_t)\,dW_t, \quad X_0 = x; \\ -dP_t = (A_t^*P_t + Q_tX_t + C_t^*Z_t)\,dt - Z_t\,dW_t, \quad P_T = MX_T; \\ N_tu_t + B_t^*P_t + D_t^*Z_t = 0. \end{cases}$$

Denote by (X^i, P^i, Z^i, u^i) solves the above forward-backward stochastic system corresponding the initial state $x = e_i$ (the unit vector in $\mathbb{R}^n$ with the ith component being 1 and others being zeroes). Define

$$\mathcal{X} := (X^1, \dots, X^n), \quad \mathcal{P} := (P^1, \dots, P^n), \quad \mathcal{U} := (\bar{u}^1, \dots, \bar{u}^n).$$

One can show that $\mathcal{P} = K\mathcal{X}$.

$$\bar{u}_t = -(N_t + D_t^*K_tD_t)^{-1}(K_tB_t + C_t^*K_tD_t)^*\bar{X}_t = F(t, K_t)\bar{X}_t.$$

$$\mathcal{U}_t = -(N_t + D_t^*K_tD_t)^{-1}(K_tB_t + C_t^*K_tD_t)^*\mathcal{X}_t = F(t, K_t)\mathcal{X}_t.$$

Optimally closed system:

$$d\mathcal{X}_t = (A_t + B_tF(t, K_t))\mathcal{X}_t\,dt + (C_t + D_tF(t, K_t))\mathcal{X}_t\,dW_t, \quad \mathcal{X}_0 = I_{n\times n}.$$

Then $K = \mathcal{P}\mathcal{X}^{-1}$.

(iii) Method of dynamic programming principle

Define

$$l(t, x, u) := \langle Q_t x, x\rangle + \langle N_t u, u\rangle.$$

$$V(s, x) = \inf_{u \in \mathscr{L}^2_{\mathscr{F}}(s,T;R^m)} E[J(u; s, x)], \quad (s, x) \in [0, T] \times R^n.$$

The value function $V(t, \cdot)$ is shown to satisfy the parallelogram, and therefore,

$$V(t, x) = \langle K_t x, x\rangle, \quad x \in R^n \quad \text{with} \quad K \in \mathscr{S}^n, \quad K \geq 0.$$

Further, K_t is absolutely continuous.

Dynamic programming principle (DPP):

$$V(s, x) = \inf_u E^u_{s,x} \left\{ \int_s^t l(r, X_r, u_r)\, dr + V(t, X_t) \right\}.$$

Bellman equation (using Itô's formula): $\forall x \in R^n$,

$$\partial_t V(t, x) + \inf_{u \in \mathbb{R}^m} \left\{ \frac{1}{2} \mathscr{L}^u_{t,x} V(t, x) + l(t, x) \right\} = 0, \quad V(T, x) = \langle Mx, x\rangle;$$

Equivalently, K is a solution of the Riccati equation.

3. The case of random coefficients

Jean-Michel Bismut's study in 1970s and BSDE

JM Bismut, Analyse convex et probabilités, Doctoral dissertation, 1973

JM Bismut (1973), Conjugate convex functions in optimal stochastic control,
J. Math Anal. Appl. 44, 384–404. linear BSDEs

JM Bismut (1976),
Linear quadratic optimal stochastic control with random coefficients
SIAM J. Control Optim. 14(3), 419–444. Quadratic BSDEs
Received December 4, 1972; Revised January 11, 1975;
published May 1976 (3.5 years later)

JM Bismut (1978),
Contrôle des systems lineares quadratiques: applications de l'intégrale stochastique,
in: Séminare de Probabilités XII, Lecture Notes in Math 649, Springer, 180–264.

Backward stochastic Riccati equation

Let (K, L) be the unique solution of the backward stochastic Riccati equation (BSRE):

$$\begin{cases} dK_t &= -[A_t^* K_t + K_t A_t + C_t^* K_t C_t + C_t^* L_t + L_t C_t + Q_t - G(t, K_t, L_t)]\, dt \\ & \quad + L_t\, dW_t, \\ K_T &= M. \end{cases}$$

The function G is nonlinear:

$$G(t, K, L) = (KB_t + C_t^* KD_t + LD_t)(N_t + D_t^* KD_t)^{-1}(KB_t + C_t^* KD_t + LD_t)^*.$$

It is a nonlinear BSDE.

The optimal control law

$$\bar{u}_t = -(N_t + D_t^* K_t D_t)^{-1}(K_t B_t + C_t^* K_t D_t + L_t D_t)^* \bar{X}_t.$$

Bismut only solved the following special case: the coefficients are independent of those underlying Brownian motions which are directly controlled, which ensures that the generator of the corresponding Riccati equation grows linearly in the second unkown variable L.

Kobylonsky (2000, AP): BSDEs and PDEs with quadratic growth

Bellman's quasi-linearization (Newton's algorithm)

Peng (1992, SICON)

$$F(t,K,L) := -(N_t + D_t^* K D_t)^{-1}(KB_t + C_t^* K D_t + LD_t)^*,$$

$$\widehat{Q}_t(K,L) := Q_t + F^*(t,K,L)N_t F(t,K,L),$$

$$\widehat{A}_t(K,L) := A + BF(t,K,L), \quad \widehat{C}_t(K,L) := C + DF(t,K,L).$$

$$dK_t = -[\widehat{A}_t^*(\cdot)K_t + K_t\widehat{A}_t(\cdot) + \widehat{C}_t^*(\cdot)K_t\widehat{C}_t(\cdot) + \widehat{C}_t^*(\cdot)L_t + L_t\widehat{C}_t(\cdot) + \widehat{Q}_t(\cdot)]\,dt + L_t\,dW_t,$$

$$K_T = M.$$

Iterating algorithm: for $i = 1, 2, \ldots,$

$$dK_t^i = -\Big[\widehat{A}_t^*(K_t^{i-1}, L_t^{i-1})K_t^i + K_t^i\widehat{A}_t(\cdot) + \widehat{C}_t^*(\cdot)K_t^i\widehat{C}_t(\cdot)$$

$$+\widehat{C}_t^*(\cdot)L_t^i + L_t^i\widehat{C}_t(\cdot) + \widehat{Q}_t(\cdot)\Big]\,dt + L_t^i\,dW_t, \qquad K_T^i = M.$$

Here arises the new issue: the convergence of L^i in the BMO space ?

4. The method of stochastic maximum principle (SMP)

Tang (2003, SICON): stochastic flow derived by the optimal closed system has an inverse.

$$\begin{cases} dX_t = (A_t X_t + B_t u_t)\, dt + (C_t X_t + D_t u_t)\, dW_t, \quad X_0 = x; \\ -dP_t = (A_t^* P_t + Q_t X_t + C_t^* Z_t)\, dt - Z_t\, dW_t, \quad P_T = M X_T; \\ N_t u_t + B_t^* P_t + D_t^* Z_t = 0. \end{cases}$$

Denote by (X^i, P^i, Z^i, u^i) solves the above forward-backward stochastic system corresponding the initial state $x = e_i$ (the unit vector in $\mathbb{R}^n$ with the ith component being 1 and others being zeroes). Define

$$\mathcal{X} := (X^1, \ldots, X^n), \quad \mathcal{P} := (P^1, \ldots, P^n), \quad \mathcal{U} := (\bar{u}^1, \ldots, \bar{u}^n).$$

Since one expect that

$$\mathcal{P} = K\mathcal{X}.$$

Then $K = \mathcal{P}\mathcal{X}^{-1}$. Define the stopping time τ_k for $k = 1, \ldots,$

$$\tau_k := \inf\left\{ t \geq 0 : |\mathcal{X}_t| \leq \frac{1}{k} \right\} \bigwedge T.$$

We show that $\lim_{k\to\infty} \tau_k = T$.

4. The method of SMP (continued)

$$\bar{u}_t = -(N_t + D_t^* K_t D_t)^{-1}(K_t B_t + C_t^* K_t D_t + L_t D_t)^* \bar{X}_t := F(t, K_t, L_t)\bar{X}_t.$$

$$\mathcal{U}_t = -(N_t + D_t^* K_t D_t)^{-1}(K_t B_t + C_t^* K_t D_t + L_t D_t)^* \mathcal{X}_t = F(t, K_t, L_t)\mathcal{X}_t.$$

Optimally closed system:

$$d\mathcal{X}_t = (A_t + B_t F(t, K_t, L_t))\mathcal{X}_t\, dt + (C_t + D_t F(t, K_t, L_t))\mathcal{X}_t\, dW_t, \quad \mathcal{X}_0 = I_{n\times n}.$$

5. The method of dynamic programming principle (DDP)

The cost functional

$$J(u;\tau,x) = \frac{1}{2}E\left\{\langle MX_T, X_T\rangle + \int_\tau^T [\langle Q_t X_t, X_t\rangle + \langle N_t u_t, u_t\rangle]\,dt \,\middle|\, \mathscr{F}_\tau\right\}.$$

The value $\mathscr{T}$-system:

$$V(\tau,x) := \operatorname*{ess.inf}_{u\in\mathscr{L}^2_{\mathscr{F}}(\tau,T;R^m)} J(u;\tau,x), \quad \tau \in \mathscr{T} \times R^n.$$

It is shown to be aggregated by an RCLL process $\mathbb{V}$:

$$V(\tau,x) = \mathbb{V}(t,x)|_{t=\tau}, a.e. \quad \forall\tau \in \mathscr{T}.$$

$$\mathbb{V}(t,x) = \langle K_t x, x\rangle. \tag{1}$$

5. The method of DPP (continued)

Bellman's Principle. Let Assumptions (A1) and (A2) be satisfied. We have
(i) For $s \le t \le T$ and $\xi \in L^2(\Omega, \mathscr{F}_s, P; \mathbb{R}^n)$,

$$\mathbb{V}(s,\xi) = \operatorname{ess.}\inf_{u\in\mathscr{U}_s} E^{s,\xi;u}\left\{\int_s^t l(r,X_r,u_r)\,dr + \mathbb{V}(t,X_t)\,\middle|\,\mathscr{F}_s\right\}.$$

For the optimal control $\overline{u} \in \mathscr{U}_s$, we have

$$\mathbb{V}(s,\xi) = E^{s,\xi;\overline{u}}\left\{\int_s^t l(r,X_r,\overline{u}_r)\,dr + \mathbb{V}(t,X_t)\,\middle|\,\mathscr{F}_s\right\}.$$

(ii) For $(s,x,u) \in [0,T]\times\mathbb{R}^n\times\mathscr{U}_s$, the process

$$\kappa_t^{s,x;u} := \mathbb{V}(t,X_t^{s,x;u}) + \int_s^t l(r,X_r^{s,x;u},u_r)\,dr$$

defined for $t\in[s,T]$, is a submartingale w.r.t. $\{\mathscr{F}_t\}$; and for the optimal control $\overline{u}\in\mathscr{U}_s$, the process $\kappa_t^{s,x;\overline{u}}, t\in[s,T]$, is a martingale w.r.t. $\{\mathscr{F}_t\}$.

We need use Itô-Wentzell's formula to derive the associated Bellman's equation. The value field $\mathbb{V}$ is a semi-martingale.

K is an essentially bounded nonnegative symmetric matrix-valued continuous semi-martingale of the form

$$K_t = K_0 - \int_0^t dk_s + \sum_{i=1}^{d} \int_0^t L_s^i \, dW_s^i, \quad t \in [0, T]; \quad K_T = M \tag{2}$$

with k being an $n \times n$ atrix-valued continuous process of bounded variation such that

$$dk_s = \ G(s, K_s, L_s) \, ds, \quad \text{almost everywhere } (s, \omega) \in [0, T] \times \Omega. \tag{3}$$

and

$$E \left[\left(\int_0^T |L_s|^2 \, ds \right)^p \right] < \infty, \quad \forall p \geq 2. \tag{4}$$

Assertion; K is a semi-martingale.

Let e_i be the unit column vector of $\mathbb{R}^n$ whose i-th component is the number 1 for $i = 1, \ldots, n$. In view of the dynamic programming principle, we see that for $x = e_i, e_i + e_j, e_i - e_j, i, j = 1, \ldots, n$, $\{\kappa_t^{0,x;0}, t \in [0,T]\}$ is a sub-martingale, and since

$$|\kappa_t^{0,x;0}| \leq \lambda |X_t^{0,x;0}|^2 + \int_0^t |X_s^{0,x;0}|^2 \, ds \leq \lambda \max_{t \in [0,T]} |X_t^{0,x;0}|^2 \in L^1(\Omega, \mathscr{F}_T, P),$$

it is of class D. Since $V(t,x)$ is continuous in the sense of conditional mean in t, $\{\kappa_t^{0,x;0}, t \in [0,T]\}$ is continuous in the sense of conditional mean in s. In view of Doob-Meyer decomposition, its bounded variational process is continuous and increasing in time, and $\{\kappa_t^{0,x;0}, t \in [0,T]\}$ is sample continuous. Define the $n \times n$ symmetric matrix-valued process

$$\Gamma_t := (\kappa_t(i,j))_{1 \leq i,j \leq n} \tag{5}$$

where

$$\kappa_t(i,i) := \kappa_t^{0,e_i;0}, \quad \kappa_t(i,j) := \frac{1}{4}[\kappa_t^{0,e_i+e_j;0} - \kappa_t^{0,e_i-e_j;0}], \quad 1 \leq i \neq j \leq n. \tag{6}$$

It is a $n \times n$ matrix-valued semi-martingale and the bounded variational process in the Doob-Meyer decomposition is continuous in time.

Define

$$\Phi_t := (X_t^{0,e_1;0}, \cdots, X_t^{0,e_n;0}), \quad t \in [0,T].$$

Then, we have

$$\Gamma_t = \Phi_t' K_t \Phi_t + \int_0^t \Phi_r' Q_r \Phi_r \, dr, \quad t \in [0,T]; \tag{7}$$

and Φ satisfies the following matrix-valued stochastic differential equation (SDE):

$$d\Phi_t = A_t \Phi_t \, dt + C_t^i \Phi_t \, dW_t^i, \quad t \in (0,T]; \quad \Phi_0 = I_n. \tag{8}$$

It is well-known that Φ_t has an inverse $\Psi_t := \Phi_t^{-1}$, satisfying the following SDE:

$$d\Psi_t = \Psi_t(-A_t + C_t^i C_t^i) \, dt - \Psi_t C_t^i \, dW_t^i, \quad t \in (0,T]; \quad \Psi_0 = I_n. \tag{9}$$

Therefore, we have

$$K_t = \Psi_t' \left(\Gamma_t - \int_0^t \Phi_r' Q_r \Phi_r \, dr \right) \Psi_t, \quad t \in [0,T]. \tag{10}$$

Since Γ is a semi-martingale, using Itô-Wentzell formula, we see that K is a semi-martingale of a form from the Doob-Meyer decomposition, with the bounded variational process k being continuous in time.

6. Unity of LQ theory for distributed parameter and s-tochastic systems: comments and perspective

R. Kalman, Contributions to the theory of optimal control (Boletin Soc. Mat. Mex., 1960)

For distributed parameter systems:

J. L. Lions (1968), Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles. Dunod.

Lukes, D. Russell (1969), The quadratic criterion for distributed systems, SICON 7, 101-121; 132-141

D. Russell (1973), The quadratic criterion in boundary laue control of linear symmetric hyperbolic systems, SICON 11, 475-509.

D. Russell (1978), Controllability and stabilizability theory for linear PADEs: recen-t progress and open questions, SIAM Review 20, 639-739

J. L. Lions (1988), SIAM Review.

I. Lasiecka, R. Triggiani (1990), Control theory for PDEs, Volm. I, II, Cambridge.

For stochastic systems:

W. Wonham (1968, SICON), On a matrix Riccati equation of stochastic control, 6, 681-697

U. G. Haussmann (1971, SICON), Optimal stationary control with state and control-dependent noise

J. M. Bismut (1973, JMAA), Conjugate convex functions in optimal stochastic control, 44, 384–404.

J. M. Bismut (1976, SICON), Linear quadratic optimal stochastic control with random coefficients, 14, 419–444.

J. M. Bismut (1978, SIAM Review): An introductory dual approach to optimal stochastic control.

S. Peng (1992, SICON), stochastic HJB equations, 30, 284-304.

S. Tang (2003, 2015, SICON): general linear quadratic optimal stochastic control with random coefficients

For stochastic distributed systems:

G. Tessitore (1992, SICON): Some remarks on the Riccati equation arising in an optimal control problem with state- and control-dependent noise, 30(3) 717–744.

G. Guatteri, G. Tessitore (2005, SICON): On the backward stochastic Riccati equation in infinite dimensions, 44(1), 159-194

G. Guatteri, G. Tessitore (2014, SICON): Well posedness of operator valued backward stochastic Riccati equations in infinite dimensional spaces, 52(6), 3776-3806

The general case remains to be solved !!

Thank you!

E-mail: sjtang@fudan.edu.cn