dr inż. Wiesław Grygierzec
Uniwersytet Rolniczy Kraków
Katedra Statystyki Matematycznej

Własności funkcji wartości dla stochastycznego sterowania optymalnego w przypadku problemu Mayera

Niech dana będzie przestrzeń probabilistyczna z filtracją $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq t_0}\}, \mathbf{P})$, na której jest zdefiniowany standardowy proces Wienera $W(t)$. Rozważamy następujący stochastyczny układ ze sterowaniem

$$\begin{cases} \mathrm{d}y(t) = f(t, y(t), \alpha(t))\, dt + \sigma(t, y(t), \alpha(t)\, dW(t), \quad t \in [t_0, \infty], \\ y(t_0) = x_0, \qquad (t_0, x_0) \in [0, \infty] \times \mathbb{R}^d \text{ ustalone}, \end{cases} \tag{1}$$

gdzie $y(t) \in \mathbb{R}^d$ jest wektorem stanu, natomiast $\alpha(t) = \alpha(t, \omega)$ jest sterowaniem w chwili t zależnym od parametru losowego $\omega \in \Omega$. Zakładamy, że dany jest pewien domknięty zbiór $M \in \mathbb{R}^d$ oraz funkcja $g(t_1, y(t_1))$ na brzegu tego zbioru $g : [t_0, \infty] \times \partial M \to \mathbb{R}$.

Jako zbiór sterowań dopuszczalnych przyjmiemy tylko takie sterowania $\alpha(t)$, które w skończonym p.n. czasie przeprowadzają wektor stanu $y(t)$ do zbioru M, tzn. $y(t_1) \in M$, gdzie $t_1 = t_1(t_0, x_0, \alpha(\cdot), \omega)$

$$t_1 := \inf\{t \geq t_0 : y(t, t_0, x_0, \alpha) \in M\}.$$

Zbiór sterowań $\alpha(t)$ oznaczmy jako

$$\begin{aligned} \mathcal{A}_{t_0,x_0} := \{\alpha : [s, \infty) \times \Omega \to A \mid \alpha \text{ mierzalne}, \\ \{\mathcal{F}_t\}_{t\geq 0}\text{-adaptowane i } t_1(t_0, x_0, \alpha(\cdot), \omega) < \infty \text{ p.n.}\}. \end{aligned} \tag{2}$$

Załóżmy, że są spełnione założenia gwarantujące istnienie rozwiązania

$$y(t) = y(t; t_0, x_0, \alpha(t))$$

problemu (1). Zdefiniujmy *funkcję wartości*

$$v(t_0, x_0) := \inf_{\alpha \in \mathcal{A}_{t_0 x_0}} \mathbb{E} g(t_1, y(t_1; t_0, x_0, \alpha)), \tag{3}$$

przyjmując konwencję $v(t_0, x_0) = \infty$, gdy $\mathcal{A}_{t_0 x_0} = \emptyset$.

Rozważane zagadnienie jest nazywane *problemem Mayera* dla sterowania optymalnego. Trajektorię y^*, dla której zachodzi $\inf_\alpha \mathbb{E} g(t_1, y(t_1)) = \mathbb{E} g(t_1, y^*(t_1))$, nazywamy *trajektorią optymalną*.

Własność 1.

(i) *$v(t, y(t))$ jest podmartyngałem, tzn. $v(\tau_1, y(\tau_1)) \leq \mathbb{E}[v(\tau_2, y(\tau_2))|\mathcal{F}_{\tau_1}]$ p.n. dla $t_0 \leq \tau_1 \leq \tau_2 \leq t_1$*

(ii) *y^*-trajektoria optymalna $\Rightarrow v(t, y^*(t))$-martyngał.*

Własność 2. *Założenie:* $W(s,\xi)$ *funkcja określona na* $\mathbb{R}^{d+1}$ *spełniająca*

(i) $W(s,\xi) = g(s,\xi)$ *dla dowolnego* $\xi \in \partial M$,

(ii) *Dla każdego rozwiązania* $y(t) = y(t, t_0, x_0, \alpha(t))$, *pole losowe* $W(t, y(t))$ *jest podmartyngałem,*

(iii) *Dla pewnej trajektorii* $\tilde{y}(t,)$, $W(t, \tilde{y}(t))$ *jest martyngałem.*

Wtedy $\tilde{y}$ *jest trajektorią optymalną, tzn.* $\tilde{y} = y^*$.

Oznaczmy $a := \sigma\sigma'$.

Własność 3. *Założenie: Funkcja* (3) $v(s,y)$ *jest klasy* C^{12} *w pewnym otoczeniu punktu* (s,y), *punktu wewnętrznego zbioru* $Q_0 = \{(s,y) : \mathcal{A}_{s,y} \neq \emptyset\}$.
Wówczas dla dowolnego $\alpha \in A$

$$\frac{\partial v}{\partial s}(s,y) + \frac{\partial v}{\partial y}(s,y) f(s,y,\alpha) + \frac{1}{2} tr[D_y^2 v(s,y) \cdot a(s,y,\alpha)] \geq 0.$$

Jeżeli α^* *jest sterowaniem optymalnym w* $\mathcal{A}_{sy}$ *to*

$$\frac{\partial v}{\partial s}(s,y) + \inf_{\alpha \in A} \left\{ \frac{\partial v}{\partial y}(s,y) f(s,y,\alpha(t)) + \frac{1}{2} tr[D_y^2 v(s,y) \cdot a(s,y,\alpha(t))] \right\} = 0. \quad (4)$$

Własność 4. *Założenie:* $w(s,y)$ *rozwiązanie klasy* C^{12} *równania* (4) *z warunkiem brzegowym* $w(s,y) = g(s,y)$ *dla* $y \in \partial M$.
Wówczas $w(s, y(s))$ *jest podmartyngałem. Jeżeli ponadto* $\tilde{\alpha}$ *jest sterowaniem w* $[t_0, t_1]$, *a* $\tilde{y}(s) = \tilde{y}(s, \tilde{\alpha}(s))$ *odpowiadającą mu trajektorią, takim, że*

$$\frac{\partial v}{\partial s}(s,\tilde{y}) + \frac{\partial v}{\partial \tilde{y}}(s,\tilde{y}) f(s,\tilde{y},\tilde{\alpha}) + \frac{1}{2} tr[D_y^2 v(s,\tilde{y}) \cdot a(s,\tilde{y},\tilde{\alpha})] = 0.$$

Wówczas $\tilde{\alpha}$ *jest sterowaniem optymalnym i* $w(s,y) = v(s,y)$ *oraz* $\tilde{y} = y^*$ *trajektorią optymalną.*