Stochastische Kontrolltheorie und Optimierung

Übungsblatt 2

Abgabe: Dienstag, 22. November, 12:00 Uhr.

Aufgabe 1

Auf der nächsten Seite dieses Übungsblatts wurde versucht das Bellman Prinzip (Theorem 1) zu beweisen. Leider sind einige der Argumente lückenhaft bzw. sogar fehlerhaft. Suchen Sie vier Lücken/Fehler im Beweis (und geben Sie genau an, worin die Lücke/der Fehler besteht).

Aufgabe 2

Auf der letzten Seite dieses Übungsblatts finden Sie eine heuristische Herleitung der sogenannten Hamilton-Jacobi-Bellman Gleichung. Welche Annahmen müssen getroffen werden, damit diese heuristische Herleitung mathematisch korrekt ist?
Hinweis: Es gibt vier Stellen in der Herleitung, an denen zusätzliche Annahmen getroffen werden müssen.

Aufgabe 3 (HJB Gleichung falls $T = \infty$)

Sei $r > 0$. Wie lautet die HJB Gleichung, wenn wir formal $T = \infty$ setzen und das Performance-Funktional $\mathcal{J}$ durch

$$\hat{\mathcal{J}}(x, \nu) \triangleq \mathrm{E}\Big[\int_0^\infty e^{-rt} f\big(X_{0,x}^\nu(t), \nu(t)\big)\mathrm{d}t\Big]$$

ersetzen?
Hinweis: Die Wertfunktion ist in diesem Fall definiert als

$$\mathcal{V}(x) \triangleq \sup_{\nu \in \mathcal{A}(0,x)} \hat{\mathcal{J}}(x, \nu), \qquad x \in \mathcal{O}.$$

Die Wertfunktion ist also nicht zeitabhängig.
Hinweis: Sie dürfen folgende Variante des Bellman Prinzips verwenden:

$$\mathcal{V}(x) = \sup_{\nu \in \mathcal{A}(0,x)} \mathrm{E}\Big[\int_0^\tau e^{-rt} f\big(X_{0,x}^\nu(t), \nu(t)\big)\mathrm{d}t + e^{-r\tau}\mathcal{V}(X_{0,x}^\nu(\tau))\Big].$$

Aufgabe 4 (HJB Gleichung im Fall von Diskontierung)

Sei $r > 0$. Wie lautet die HJB Gleichung, wenn wir das Performance-Funktional $\mathcal{J}$ durch

$$\tilde{\mathcal{J}}(t, x, \nu) \triangleq \mathrm{E}\Big[\int_t^T e^{-r(s-t)} f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + e^{-r(T-t)} g(X_{t,x}^\nu(T))\Big]$$

ersetzen?
Hinweis: Raten Sie zunächst, wie das Bellman Prinzip in diesem Fall aussieht (z.B. anhand vom Beweis von Theorem 1). Leiten Sie daraus heuristisch die HJB Gleichung her.

Theorem 1 (Bellman Prinzip)
Seien $(t,x) \in \overline{\mathcal{T}} \times \mathcal{O}$ und sei τ eine $\overline{\mathcal{T}}_t$-wertige Stoppzeit. Dann gilt das Bellman Prinzip:

$$\mathcal{V}(t,x) = \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big].$$

Beweisversuch. Schritt 1: Wir zeigen, dass

$$\mathcal{V}(t,x) \leq \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big].$$

Sei dafür $\nu \in \mathcal{A}(t,x)$ beliebig gewählt. Dann gilt

$$\begin{aligned}
\mathcal{J}(t,x,\nu) &= \mathrm{E}\Big[\int_t^T f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + g(X_{t,x}^\nu(T))\Big] \\
&= \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathrm{E}\Big[\int_\tau^T f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + g(X_{t,x}^\nu(T))\Big|\mathfrak{F}(\tau)\Big]\Big] \\
&= \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{J}\big(\tau, X_{t,x}^\nu(\tau), \nu\big)\Big] \\
&\leq \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big].
\end{aligned}$$

Da dies für jedes $\nu \in \mathcal{A}(t,x)$ gilt, folgt

$$\mathcal{V}(t,x) = \sup_{\nu \in \mathcal{A}(t,x)} \mathcal{J}(t,x,\nu) \leq \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big].$$

Schritt 2: Wir zeigen, dass

$$\mathcal{V}(t,x) \geq \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big].$$

Für $\varepsilon > 0$ und $\nu \in \mathcal{A}(t,x)$ gegeben, wähle einen ε-optimalen Kontrollprozess $\nu_\varepsilon \in \mathcal{A}(\tau, X_{t,x}^\nu(\tau))$, d.h. ν_ε erfüllt

$$\mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big) \leq \mathcal{J}\big(\tau, X_{t,x}^\nu(\tau), \nu_\varepsilon\big) + \varepsilon.$$

Definiere nun einen Kontrollprozess $\hat{\nu} = \{\hat{\nu}(s)\}_{s \in \overline{\mathcal{T}}_t}$ durch

$$\hat{\nu}(s) \triangleq \nu(s)\mathbb{1}_{[t,\tau]}(s) + \nu_\varepsilon(s)\mathbb{1}_{(\tau,T]}(s) \qquad \text{für alle } s \in \overline{\mathcal{T}}.$$

Dann gilt wie in Schritt 1

$$\begin{aligned}
\mathcal{V}(t,x) \geq \mathcal{J}(t,x,\hat{\nu}) &= \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{J}\big(\tau, X_{t,x}^\nu(\tau), \nu_\varepsilon\big)\Big] \\
&\geq \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big] - \varepsilon.
\end{aligned}$$

Da ν und ε beliebig gewählt waren, folgt

$$\mathcal{V}(t,x) \geq \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^\tau f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(\tau, X_{t,x}^\nu(\tau)\big)\Big]. \qquad \square$$

Heuristische Herleitung der Hamilton-Jacobi-Bellman Gleichung

Wir leiten aus dem Bellman Prinzip heuristisch eine partielle Differentialgleichung für die Wertfunktion $\mathcal{V}$ her. Seien dafür $t \in \mathcal{T}$ und $x \in \mathcal{O}$. Für ein $\varepsilon > 0$ mit $t + \varepsilon < T$ wählen wir dafür die Stoppzeit im Bellman Prinzip als $\tau \triangleq t + \varepsilon$ und erhalten

$$\mathcal{V}(t,x) = \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^{t+\varepsilon} f\big(X_{t,x}^\nu(s), \nu(s)\big)\mathrm{d}s + \mathcal{V}\big(t+\varepsilon, X_{t,x}^\nu(t+\varepsilon)\big)\Big]. \tag{1}$$

Wir wenden nun die Itô-Formel auf $\mathcal{V}(t+\varepsilon, X_{t,x}^\nu(t+\varepsilon))$ an und erhalten

$$\begin{aligned}\mathcal{V}\big(t+\varepsilon, X_{t,x}^\nu(t+\varepsilon)\big) &= V(t,x) + M_{t,x}^\nu(t+\varepsilon) \\ &\qquad + \int_t^{t+\varepsilon}\big[\partial_t \mathcal{V}\big(s, X_{t,x}^\nu(s)\big) + \mathcal{L}_{\nu(s)}\mathcal{V}\big(s, X_{t,x}^\nu(s)\big)\big]\mathrm{d}s,\end{aligned} \tag{2}$$

wobei der Prozess $M_{t,x}^\nu = \{M_{t,x}^\nu(s)\}_{s \in \overline{\mathcal{T}}_t}$ gegeben ist durch

$$M_{t,x}^\nu(s) \triangleq \int_t^s \mathrm{D}_x \mathcal{V}\big(r, X_{t,x}^\nu(r)\big)\sigma\big(X_{t,x}^\nu(r), \nu(r)\big)\mathrm{d}W(r), \qquad s \in \overline{\mathcal{T}}_t,$$

und der Differentialoperator $\mathcal{L}_u$ für $u \in \mathcal{U}$ definiert ist durch

$$\mathcal{L}_u \mathcal{V}(t,x) \triangleq b(x,u)^\top \mathrm{D}_x \mathcal{V}(t,x) + \tfrac{1}{2}\operatorname{tr}\big[\sigma(x,u)\sigma(x,u)^\top \mathrm{D}_x^2 \mathcal{V}(t,x)\big].$$

Durch Einsetzen von (2) in (1) folgt dann

$$0 = \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\int_t^{t+\varepsilon}\big[\partial_t \mathcal{V}\big(s, X_{t,x}^\nu(s)\big) + \mathcal{L}_{\nu(s)}\mathcal{V}\big(s, X_{t,x}^\nu(s)\big) + f\big(X_{t,x}^\nu(s), \nu(s)\big)\big]\mathrm{d}s\Big].$$

Wir teilen nun beide Seiten durch ε und schicken dann $\varepsilon \downarrow 0$. Aus dem Mittelwertsatz folgt dann

$$\begin{aligned}0 &= \lim_{\varepsilon \downarrow 0} \sup_{\nu \in \mathcal{A}(t,x)} \mathrm{E}\Big[\tfrac{1}{\varepsilon}\int_t^{t+\varepsilon}\big[\partial_t \mathcal{V}\big(s, X_{t,x}^\nu(s)\big) + \mathcal{L}_{\nu(s)}\mathcal{V}\big(s, X_{t,x}^\nu(s)\big) + f\big(X_{t,x}^\nu(s), \nu(s)\big)\big]\mathrm{d}s\Big] \\ &= \sup_{\nu \in \mathcal{A}(t,x)} \big[\partial_t \mathcal{V}(t,x) + \mathcal{L}_{\nu(t)}\mathcal{V}(t,x) + f(x,\nu(t))\big] \\ &= \partial_t \mathcal{V}(t,x) + \sup_{u \in \mathcal{U}}\big[\mathcal{L}_u \mathcal{V}(t,x) + f(x,u)\big].\end{aligned}$$

Wir haben also (heuristisch) argumentiert, dass die Wertfunktion $\mathcal{V}$ die partielle Differentialgleichung

$$\partial_t \mathcal{V}(t,x) + \sup_{u \in \mathcal{U}}\big[\mathcal{L}_u \mathcal{V}(t,x) + f(x,u)\big] = 0 \qquad \text{für alle } (t,x) \in \mathcal{T} \times \mathcal{O}$$

löst. Wir bezeichnen diese Gleichung als die Hamilton-Jacobi-Bellman Gleichung. Die Randbedingung für $\mathcal{V}$ ist offensichtlich

$$\mathcal{V}(T,x) = g(x) \qquad \text{für alle } x \in \mathcal{O}.$$