Corrigendum for: Rushikesh Kamalapurkar et al. *Reinforcement learning for optimal feedback control: A Lyapunov-based approach*. Communications and Control Engineering. Springer International Publishing, 2018. DOI: 10.1007/978-3-319-78384-0

Rushikesh Kamalapurkar

October 31, 2019

1. Page 7: On the line after Equation 1.14, $V(0) = 0$ should be $V^*(0) = 0$.
2. Page 7: Theorem 1.5 is incomplete as stated. Positive definiteness of V^* is also required for the Hamilton–Jacobi–Bellman equation to be necessary and sufficient for optimality. See Appendix A.

A Proof of the claim in item 2

For a controlled dynamical system described by the initial value problem

$$\dot{x} = f(x, u, t),\ x(t_0) = x_0, \tag{1}$$

where t_0 is the initial time, $x \in \mathbb{R}^n$ denotes the system state, $u \in U \subset \mathbb{R}^m$ denotes the control input, and U denotes the action-space, consider a family (parameterized by t) of optimal control problems described by the cost functionals

$$J(t, y, u(\cdot)) = \int_t^\infty L(\phi(\tau; t, y, u(\cdot)), u(\tau), \tau)\, \mathrm{d}\tau \tag{2}$$

where $L : \mathbb{R}^n \times U \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is the **Lagrange cost, with** $L(x, u, t) \geq 0$**, for all** $(x, u, t) \in \mathbb{R}^n \times U \times \mathbb{R}_{\geq 0}$, and the notation $\phi(\tau; t, y, u(\cdot))$ is used to denote a trajectory of the system in (1), evaluated at time τ, under the controller $u : \mathbb{R}_{\geq t_0} \to U$, starting at the initial time t, and with the initial state y. The short notation $x(\tau)$ is used to denote $\phi(\tau; t, y, u(\cdot))$ when the controller, the initial time, and the initial state are clear from the context. Throughout this discussion, it is assumed that the controllers and the dynamical systems are such that the initial value problem in (1) admits a unique complete solution starting from any initial condition.

Let the optimal value function $V^* : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ be defined as

$$V^*(x, t) \coloneqq \inf_{u_{[t,\infty)} \in \mathcal{U}_{(t,x)}} J(t, x, u(\cdot)), \tag{3}$$

where the notation $u_{[t,\infty)}$ for $t \geq t_0$ denotes the controller $u(\cdot)$ restricted to the time interval $[t, \infty)$ and $\mathcal{U}_{(t,x)}$ denotes the set of controllers that are admissible for x. The definition of admissibility is weaker in the first theorem, but needs to be strengthened for the other theorems.

The following theorem is a generalization of Theorem 1.2 from the book to infinite horizon problems.

Theorem 1. *Given $t_0 \in \mathbb{R}_{\geq 0}$, $x_0 \in \mathbb{R}^n$, let the class of admissible controllers, $\mathcal{U}_{(t_0,x_0)}$, include all Lebesgue measurable locally bounded controllers so that the initial value problem in* (1) *admits a unique complete solution starting from (t_0, x_0). Assume that the optimal value function is continuously differentiable, i.e., $V^* \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}_{\geq t_0}, \mathbb{R})$. If there exists a function $V : \mathbb{R}^n \times \mathbb{R}_{\geq t_0} \to \mathbb{R}$ such that*

1. *$V \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}_{\geq t_0}, \mathbb{R})$ and V satisfies the Hamilton-Jacobi-Bellman equation*

$$0 = -\nabla_t V(x, t) - \inf_{\mu \in U} \left\{ L(x, \mu, t) + \nabla_x V^T(x, t) f(x, \mu, t) \right\}, \tag{4}$$

   for all $t \in [t_0, \infty)$ and all $x \in \mathbb{R}^n$,

2. *for every controller $v(\cdot) \in \mathcal{U}_{(t_0,x_0)}$ for which there exists $M_v \geq 0$ so that $\int_{t_0}^t L(\phi(\tau, t_0, x_0, v(\cdot)), v(\tau), \tau)\, \mathrm{d}\tau \leq M_v$ for all $t \in \mathbb{R}_{\geq t_0}$, the function V, evaluated along the resulting trajectory, satisfies*

$$\lim_{t \to \infty} V(\phi(t; t_0, x_0, v(\cdot))) = 0, \tag{5}$$

   and

3. *there exists $u(\cdot) \in \mathcal{U}_{(t_0,x_0)}$, such that the function V, the controller $u(\cdot)$, and the trajectory $x(\cdot)$ of* (1) *under $u(\cdot)$ with the initial condition $x(t_0) = x_0$, satisfy, the Hamiltonian minimization condition*

$$\begin{aligned} &L(x(t), u(t), t) + \nabla_x V^T(x(t), t) f(x(t), u(t), t) \\ &\qquad = \min_{\mu \in U} \left\{ L(x(t), \mu, t) + \nabla_x V^T(x(t), t) f(x(t), \mu, t) \right\}, \quad \forall t \in \mathbb{R}_{\geq t_0}, \end{aligned} \tag{6}$$

   and the bounded cost condition

$$\exists M_u \geq 0 \quad | \quad \int_{t_0}^t L(x(\tau), v(\tau), \tau)\, \mathrm{d}\tau \leq M_u, \quad \forall t \in \mathbb{R}_{\geq t_0}, \tag{7}$$

then, $V(t_0, x_0)$ is the optimal cost (i.e., $V(t_0, x_0) = V^*(t_0, x_0)$) and $u(\cdot)$ is an optimal controller.

Furthermore, if $V^* \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ is the optimal value function, then it satisfies the HJB equation in (4).

Proof. Let $x(\cdot) \coloneqq \phi(\cdot; t_0, x_0, u(\cdot))$, where $u(\cdot)$ is an admissible controller that satisfies (6) and (7), and $y(\cdot) \coloneqq \phi(\cdot; t_0, x_0, v(\cdot))$ where $v(\cdot)$ is any other admissible controller. The Hamiltonian minimization condition in (6) implies that along the trajectory $x(\cdot)$, the control $\mu = u(t)$ achieves the infimum in (4) for all $t \in \mathbb{R}_{\geq t_0}$. Thus, along the trajectory $x(\cdot)$, (4) implies that

$$-\nabla_t V(x(t), t) - \nabla_x V^T(x(t), t) f(x(t), u(t), t) = L(x(t), u(t), t)$$

That is,

$$-\frac{\mathrm{d}}{\mathrm{d}t} V(x(t), t) = L(x(t), u(t), t). \tag{8}$$

Since V satisfies the HJB equation everywhere, it is clear that along the trajectory $y(\cdot)$,

$$\inf_{\mu \in U} \left\{ L(y(t), \mu, t) + \nabla_x V^T(y(t), t) f(y(t), \mu, t) \right\} \leq L(y(t), v(t), t) + \nabla_x V^T(y(t), t) f(y(t), v(t), t)$$

and as a result, the HJB equation, evaluated along $y(\cdot)$, yields

$$0 \geq -\nabla_t V(y(t), t) - \nabla_x V^T(y(t), t) f(y(t), v(t), t) - L(y(t), v(t), t).$$

That is,

$$-\frac{\mathrm{d}}{\mathrm{d}t} V(y(t), t) \leq L(y(t), v(t), t). \tag{9}$$

Integrating (8) and (9) over a finite interval $[t_0, T]$,

$$-\int_{t_0}^{T} \frac{\mathrm{d}}{\mathrm{d}t} V(x(t), t)\, \mathrm{d}t = (V(x(t_0), t_0) - V(x(T), T)) = \int_{t_0}^{T} L(x(t), u(t), t)\, \mathrm{d}t,$$

and

$$-\int_{t_0}^{T} \frac{\mathrm{d}}{\mathrm{d}t} V(y(t), t)\, \mathrm{d}t = (V(y(t_0), t_0) - V(y(T), T)) \leq \int_{t_0}^{T} L(y(t), v(t), t)\, \mathrm{d}t.$$

Since $x(t_0) = y(t_0) = x_0$, it can be concluded that

$$V(x_0, t_0) = \int_{t_0}^{T} L(x(t), u(t), t)\, \mathrm{d}t + V(x(T), T) \leq \int_{t_0}^{T} L(y(t), v(t), t)\, \mathrm{d}t + V(y(T), T), \forall T \in \mathbb{R}_{\geq t_0},$$

and as a result,

$$V(x_0, t_0) = \lim_{T \to \infty} \int_{t_0}^{T} L(x(t), u(t), t)\, \mathrm{d}t + V(x(T), T) \leq \lim_{T \to \infty} \int_{t_0}^{T} L(y(t), v(t), t)\, \mathrm{d}t + V(y(T), T).$$

Since $u(\cdot)$ satisfies (7) and $(x, u, t) \mapsto L(x, u, t)$ is nonnegative, the improper integral $\int_{t_0}^{\infty} L(x(t), u(t), t)\, \mathrm{d}t$ exists, is bounded, and equal to the total cost $J(t_0, x_0, u(\cdot))$. Taking (5) into account, it can thus be concluded that

$$\lim_{T \to \infty} \int_{t_0}^{T} L(x(t), u(t), t)\, \mathrm{d}t + V(x(T), T) = J(t_0, x_0, u(\cdot))$$

If $v(\cdot)$ satisfies (7), then a similar analysis yields

$$\lim_{T \to \infty} \int_{t_0}^{T} L(y(t), v(t), t)\, \mathrm{d}t + V(y(T), T) = J(t_0, x_0, v(\cdot)),$$

and as a result,

$$V(x_0, t_0) = J(t_0, x_0, u(\cdot)) \leq J(t_0, x_0, v(\cdot)). \tag{10}$$

If $v(\cdot)$ does not satisfy (7), then nonnegativity of $(x, u, t) \mapsto L(x, u, t)$ implies that the total cost resulting from $v(\cdot)$ is unbounded and (10) holds canonically. In conclusion, $V(t_0, x_0)$ is the optimal cost (i.e., $V(t_0, x_0) = V^*(t_0, x_0)$) and $u(\cdot)$ is an optimal controller.

Under the assumptions made herein, proof of the fact that optimal value functions for infinite horizon problems satisfy HJB equations is identical to the proof of that fact for finite horizon problems (see Theorem 1.2 in the book). □

For the next theorem, a controller $v : \mathbb{R}_{\geq t_0} \to U$ is said to be admissible for a given initial state (t_0, x_0) if it is bounded, generates a unique bounded trajectory starting from x_0, and results in bounded total cost. An admissible controller that results in the smallest cost among all admissible controllers will be called an optimal controller. In the following result, since the dynamics and the Lagrange cost are time-invariant, if $v : \mathbb{R}_{\geq t_0} \to U$ is admissible for a given initial state (t_0, x_0), then $v' : \mathbb{R}_{\geq t_1} \to U$, defined as $v'(t) = v(t + t_0 - t_1)$, for all $t \in \mathbb{R}_{\geq t_1}$ is admissible for (t_1, x_0), and trajectories of the system starting from (t_0, x_0) under $v(\cdot)$ and those starting from (t_1, x_0) under $v'(\cdot)$ are identical. As a result, the set of admissible controllers, system trajectories, value functions, and total costs can be considered independent of t_0 without loss of generality. The following two theorems prove the claim in item 2. The proofs are detailed in order to demonstrate the need for admissibility restrictions for some statements, and to ensure validity of some other statements under the admissibility restrictions.

Theorem 2. *Consider the optimal control problem*

$$P: \qquad \min_{u(\cdot)\in\mathcal{U}_{x_0}} \quad J(x_0, u(\cdot)) \coloneqq \int_{t_0}^{\infty} r(\phi(\tau; x_0, u(\cdot)), u(\tau))\, \mathrm{d}\tau$$
$$\text{subject to} \quad \dot{x} = f(x) + g(x)\, u$$

where the local cost $r : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is defined as $r(x, u) \coloneqq Q(x) + u^T R u$, with $Q : \mathbb{R}^n \to \mathbb{R}$, a ***continuously differentiable*** *positive definite function and $R \in \mathbb{R}^{m\times m}$, a symmetric positive definite matrix. Assume further that the optimal value function $V^* : \mathbb{R}^n \to \mathbb{R}$ corresponding to P is continuously differentiable.*

If $x \mapsto V(x)$ is positive definite and satisfies the closed-loop Hamilton-Jacobi-Bellman equation

$$r(x, \psi(x)) + \nabla_x V(x)\,(f(x) + g(x)\,\psi(x)) = 0, \qquad \forall x \in \mathbb{R}^n, \tag{11}$$

with

$$\psi(x) = -\frac{1}{2} R^{-1} g^T(x)\,(\nabla_x V(x))^{\mathrm{T}}, \tag{12}$$

then $V(\cdot)$ is the optimal value function and the the state-feedback law $u(t) = \psi(x(t))$ is the optimal controller.

Proof. Note that (11), along with positive definiteness of Q, R, and V, imply that under the state-feedback law $u(t) = \psi(x(t))$, the closed-loop system $\dot{x} = f(x) + g(x)\,\psi(x)$ is globally asymptotically stable. Furthermore, since $V(0) = 0$, every trajectory of the closed-loop system converges to the origin and since (11) holds for all $x \in \mathbb{R}^n$, and in particular, holds along every trajectory of the closed-loop system, it can be concluded that

$$\int_t^{\infty} r(x(\tau), \psi(x(\tau)))\, \mathrm{d}t = V(x(t)) = J(x(t), \psi(x(\cdot))), \forall t \in \mathbb{R}$$

along every trajectory of the closed-loop system. As a result, all control signals resulting from the state-feedback law $u(t) = \psi(x(t))$ are admissible for all initial conditions.

For each $x \in \mathbb{R}^n$ we have

$$\frac{\partial\,(r(x, u) + \nabla_x V(x)\,(f(x) + g(x)\,u))}{\partial u} = 2u^T R + \nabla_x V(x)\, g(x).$$

hence, $u = -\frac{1}{2} R^{-1} g^T(x)\,(\nabla_x V(x))^T = \psi(x)$ extremizes $r(x, u) + \nabla_x V(x)\,(f(x) + g(x)\,u)$. Furthermore, the Hessian

$$\frac{\partial^2\,(r(x, u) + \nabla_x V(x)\,(f(x) + g(x)\,u))}{\partial^2 u} = 2R$$

is positive definite. Hence, $u = \psi(x)$ minimizes $u \mapsto r(x, u) + \nabla_x V(x)\,(f(x) + g(x)\,u)$ for each $x \in \mathbb{R}^n$.

As a result, the closed-loop HJB equation (11), along with the control law (12) are equivalent to the HJB equation (4). Furthermore, all trajectories starting from all initial conditions in response to the controller $u(t) = \psi(x(t))$ satisfy the Hamiltonian minimization condition (6) and the bounded cost condition (7).

Also, given any initial condition x_0 and a controller $v(\cdot)$ that is admissible for (x_0), boundedness of the controller $v(\cdot)$ and the resulting trajectory $\phi(\cdot; t_0, x_0, v(\cdot))$, continuity of $x \mapsto f(x, u)$ and $x \mapsto g(x, u)$, and continuity of $x \mapsto \nabla_x Q(x)$ can be used to conclude that

$$t \mapsto Q\left(\phi\left(t; t_0, x_0, v\left(\cdot\right)\right)\right) \text{ is uniformly continuous.}$$

Admissibility of $v\left(\cdot\right)$ and positive definiteness of R imply that $\left|\int_{t_0}^{T} Q\left(\phi\left(t; t_0, x_0, v\left(\cdot\right)\right)\right) \mathrm{d}t\right| \leq M$ for all $T \geq t_0$ and some $M \geq 0$. Furthermore, positive definiteness of $x \mapsto Q\left(x\right)$ implies monotonicity of $T \mapsto \int_{t_0}^{T} Q\left(\phi\left(t; t_0, x_0, v\left(\cdot\right)\right)\right) \mathrm{d}t$. As a result, the limit

$$\lim_{T\to\infty} \int_{t_0}^{T} Q\left(\phi\left(t; t_0, x_0, v\left(\cdot\right)\right)\right) \mathrm{d}t \text{ exists and is bounded.}$$

By Barbalat's lemma, $\lim_{t\to\infty} Q\left(\phi\left(t; t_0, x_0, v\left(\cdot\right)\right)\right) = 0$, which, due to positive definiteness and continuity of $x \mapsto Q\left(x\right)$ implies that $\lim_{t\to\infty} \phi\left(t; t_0, x_0, v\left(\cdot\right)\right) = 0$, and finally, from continuity and positive definiteness of V,

$$\lim_{t\to\infty} V\left(\phi\left(t; t_0, x_0, v\left(\cdot\right)\right)\right) = 0,$$

which establishes (5).

Arguments similar to the proof of Theorem 1 can then be invoked to conclude that $V\left(x_0\right)$ is the optimal cost and $u\left(t\right) = \psi\left(x\left(t\right)\right)$ is the unique optimal controller among all admissible controllers. Since the initial condition was arbitrary, the proof of Theorem 2 is complete. □

To facilitate the following discussion, let $\mathcal{U}_{x,[t_1,t_2]}$ denote the space of controllers that are restrictions over $[t_1, t_2]$ of controllers admissible for x. The task is then to show that value functions satisfy HJB equations.

Theorem 3. *Consider the optimal control problem P stated in Theorem 2 and assume that for every initial condition x_0, an optimal controller that is admissible for x_0 exists. If the optimal value function corresponding to P, defined as*

$$V^*\left(x\right) \coloneqq \inf_{u(\cdot)\in\mathcal{U}_x} \int_t^{\infty} r\left(\phi\left(\tau; t, x, u\left(\cdot\right)\right), u\left(\tau\right)\right) \mathrm{d}\tau, \tag{13}$$

is continuously differentiable then it satisfies the HJB equation

$$r\left(x, \psi\left(x\right)\right) + \nabla_x V^*\left(x\right)\left(f\left(x\right) + g\left(x\right)\psi^*\left(x\right)\right) = 0, \qquad \forall x \in \mathbb{R}^n, \tag{14}$$

with

$$\psi^*\left(x\right) = -\frac{1}{2} R^{-1} g^T\left(x\right)\left(\nabla_x V^*\left(x\right)\right)^{\mathrm{T}}. \tag{15}$$

Proof. First, it will be shown that the value function satisfies the principle of optimality. To facilitate the discussion, given $x \in \mathbb{R}^n$, let $v^*_{(x,t)} : \mathbb{R}_{\geq t} \to U$ denote an optimal controller starting from the initial state x and initial time t.

Claim 1 (Principle of optimality under admissibility restrictions)**.** For all $x \in \mathbb{R}^n$, and for all $\Delta t > 0$,

$$V^*\left(x\right) = \inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \left\{ \int_t^{t+\Delta t} r\left(\phi\left(\tau; t, x, u\left(\cdot\right)\right), u\left(\tau\right)\right) \mathrm{d}\tau + V^*\left(x\left(t+\Delta t\right)\right) \right\}. \tag{16}$$

Proof of claim: Consider the function $V : \mathbb{R}^n \to \mathbb{R}$ defined as

$$V\left(x\right) \coloneqq \inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \left\{ \int_t^{t+\Delta t} r\left(\phi\left(\tau; t, x, u\left(\cdot\right)\right), u\left(\tau\right)\right) \mathrm{d}\tau + V^*\left(x\left(t+\Delta t\right)\right) \right\}$$

Based on the definition in (13)

$$V\left(x\right) = \inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \left\{ \int_t^{t+\Delta t} r\left(\phi\left(\tau; t, x, u\left(\cdot\right)\right), u\left(\tau\right)\right) \mathrm{d}\tau + \inf_{v(\cdot)\in\mathcal{U}_{x(t+\Delta t)}} \int_{t+\Delta t}^{\infty} r\left(\phi\left(\tau; t, x\left(t+\Delta t\right), v\left(\cdot\right)\right), v\left(\tau\right)\right) \mathrm{d}\tau \right\}.$$

Since the first integral is independent of the control over $\mathbb{R}_{\geq t+\Delta t}$,

$$V\left(x\right) = \inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \inf_{v(\cdot)\in\mathcal{U}_{x(t+\Delta t)}} \left\{ \int_t^{t+\Delta t} r\left(\phi\left(\tau; t, x, u\left(\cdot\right)\right), u\left(\tau\right)\right) \mathrm{d}\tau + \int_{t+\Delta t}^{\infty} r\left(\phi\left(\tau; t, x\left(t+\Delta t\right), v\left(\cdot\right)\right), v\left(\tau\right)\right) \mathrm{d}\tau \right\}.$$

Combining the integrals and using the fact that concatenation of admissible restrictions and admissible controllers result in admissible controllers, $\inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \inf_{v(\cdot)\in\mathcal{U}_{x(t+\Delta t)}}$ is equivalent to $\inf_{w(\cdot)\in\mathcal{U}_x}$, where $w:\mathbb{R}_{\geq t}\to U$ is defined as $w(\tau) \coloneqq \begin{cases} u(\tau) & t\leq\tau\leq t+\Delta t, \\ v(\tau) & \tau > t+\Delta t, \end{cases}$ it can be concluded that

$$V(x) = \inf_{w(\cdot)\in\mathcal{U}_x} \int_t^\infty r\left(\phi\left(\tau;t,x,w(\cdot)\right),w(\tau)\right)\,\mathrm{d}\tau = V^*(x).$$

Thus,

$$V(x) \geq V^*(x). \tag{17}$$

On the other hand, by the definition of the infimum, for all $\epsilon>0$, there exists a controller $u_\epsilon(\cdot)$ such that

$$V^*(x) + \epsilon \geq J\left(x,u_\epsilon(\cdot)\right).$$

Let $x_\epsilon:\mathbb{R}_{\geq t_0}\to\mathbb{R}^n$ denote the trajectory corresponding to $u_\epsilon(\cdot)$. Since the restriction $u_{\epsilon,\mathbb{R}_{\geq t_1}}(\cdot)$ of $u_\epsilon(\cdot)$ to $\mathbb{R}_{\geq t_1}$ is admissible for $x_\epsilon(t_1)$ for all $t_1>t_0$,

$$\begin{aligned} J\left(x,u_\epsilon(\cdot)\right) &= \int_t^{t+\Delta t} r\left(x_\epsilon(\tau),u_\epsilon(\tau)\right)\,\mathrm{d}\tau + J\left(x_\epsilon(t+\Delta t),u_{\epsilon,\mathbb{R}_{\geq t+\Delta t}}(\cdot)\right), \\ &\geq \int_t^{t+\Delta t} r\left(x_\epsilon(\tau),u_\epsilon(\tau)\right)\,\mathrm{d}\tau + V^*\left(x_\epsilon(t+\Delta t)\right) \geq V(x). \end{aligned}$$

Thus, $V(x)\leq V^*(x)$, which, along with (17), implies $V(x)=V^*(x)$. □

Since $V^*\in\mathcal{C}^1(\mathbb{R}^n,\mathbb{R})$, given any admissible $u(\cdot)$ and corresponding trajectory $x(\cdot)$,

$$V^*\left(x(t+\Delta t)\right) = V^*(x) + \nabla_x V^*(x)\left(\left(f(x)+g(x)u(t)\right)\Delta t\right) + o(\Delta t).$$

Furthermore,

$$\int_t^{t+\Delta t} r\left(x(\tau),u(\tau)\right)\,\mathrm{d}\tau = r\left(x,u(t)\right)\Delta t + o(\Delta t)$$

From the principle of optimality in (16),

$$V^*(x) = \inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \left\{r\left(x,u(t)\right)\Delta t + V^*(x) + \nabla_x V^*(x)\left(\left(f(x)+g(x)u(t)\right)\Delta t\right) + o(\Delta t)\right\}.$$

That is,

$$0 = \inf_{u(\cdot)\in\mathcal{U}_{x,[t,t+\Delta t]}} \left\{r\left(x,u(t)\right)\Delta t + \nabla_x V^*(x)\left(\left(f(x)+g(x)u(t)\right)\Delta t\right) + o(\Delta t)\right\}$$

Dividing by Δt and taking the limit as Δt goes to zero,

$$0 = \inf_{u\in U}\left\{r(x,u) + \nabla_x V^*(x)\left(f(x)+g(x)u\right)\right\}, \quad \forall x\in\mathbb{R}^n. \qquad \square$$

In conclusion, under the assumptions made in this document, the optimal value function is continuously differentiable, positive definite, and satisfies the HJB equation, all functions that are continuously differentiable and positive definite and satisfy the HJB equation are optimal value functions, and optimal value functions are, by definition, unique. As a result, if there is a continuously differentiable and positive definite solution of the HJB equation then it is unique and is also the optimal value function.