Linear-Quadratic-Gaussian (LQG) Controllers

Emo Todorov

Applied Mathematics and Computer Science & Engineering

University of Washington

Spring 2014

LQG in continuous time

Recall that for problems with dynamics and cost

$$\begin{aligned} d\mathbf{x} &= (\mathbf{a}(\mathbf{x}) + B(\mathbf{x})\mathbf{u})\,dt + C(\mathbf{x})\,d\boldsymbol{\omega} \\ \ell(\mathbf{x},\mathbf{u}) &= q(\mathbf{x}) + \frac{1}{2}\mathbf{u}^\mathsf{T} R(\mathbf{x})\,\mathbf{u} \end{aligned}$$

the optimal control law is $\mathbf{u}^* = -R^{-1}B^\mathsf{T} v_\mathbf{x}$ and the HJB equation is

$$-v_t = q + \mathbf{a}^\mathsf{T} v_\mathbf{x} + \frac{1}{2}\operatorname{tr}\left(CC^\mathsf{T} v_{\mathbf{xx}}\right) - \frac{1}{2} v_\mathbf{x}^\mathsf{T} B R^{-1} B^\mathsf{T} v_\mathbf{x}$$

We now impose further restrictions (LQG system):

$$\begin{aligned} d\mathbf{x} &= (A\mathbf{x} + B\mathbf{u})\,dt + C d\boldsymbol{\omega} \\ \ell(\mathbf{x},\mathbf{u}) &= \frac{1}{2}\mathbf{x}^\mathsf{T} Q\mathbf{x} + \frac{1}{2}\mathbf{u}^\mathsf{T} R\mathbf{u} \\ q_\mathcal{T}(\mathbf{x}) &= \frac{1}{2}\mathbf{x}^\mathsf{T} Q_\mathcal{T}\mathbf{x} \end{aligned}$$

Continuous-time Riccati equations

Substituting the LQG dynamics and cost in the HJB equation yields

$$-v_t = \frac{1}{2}\mathbf{x}^\mathsf{T} Q\mathbf{x} + \mathbf{x}^\mathsf{T} A^\mathsf{T} v_\mathbf{x} + \frac{1}{2}\operatorname{tr}\left(CC^\mathsf{T} v_{\mathbf{xx}}\right) - \frac{1}{2} v_\mathbf{x}^\mathsf{T} BR^{-1}B^\mathsf{T} v_\mathbf{x}$$

We can now show that v is quadratic:

$$v(\mathbf{x}, t) = \frac{1}{2}\mathbf{x}^\mathsf{T} V(t)\,\mathbf{x} + \alpha(t)$$

At the final time this holds with $\alpha(T) = 0$ and $V(T) = Q_\mathcal{T}$. Then

$$-\dot{\alpha} - \frac{1}{2}\mathbf{x}^\mathsf{T}\dot{V}\mathbf{x} = \frac{1}{2}\mathbf{x}^\mathsf{T} Q\mathbf{x} + \mathbf{x}^\mathsf{T} A^\mathsf{T} V\mathbf{x} + \frac{1}{2}\operatorname{tr}\left(CC^\mathsf{T} V\right) - \frac{1}{2}\mathbf{x}^\mathsf{T} VBR^{-1}B^\mathsf{T} V\mathbf{x}$$

Using the fact that $\mathbf{x}^\mathsf{T} A^\mathsf{T} V\mathbf{x} = \mathbf{x}^\mathsf{T} VA\mathbf{x}$ and matching powers of $\mathbf{x}$ yields

Theorem (Riccati equation)

$$\begin{aligned} -\dot{V} &= Q + A^\mathsf{T} V + VA - VBR^{-1}B^\mathsf{T} V \\ -\dot{\alpha} &= \frac{1}{2}\operatorname{tr}\left(CC^\mathsf{T} V\right) \end{aligned}$$

Linear feedback control law

When $v(\mathbf{x}, t) = \frac{1}{2}\mathbf{x}^\mathsf{T} V(t)\,\mathbf{x} + \alpha(t)$, the optimal control $\mathbf{u}^* = -R^{-1}B^\mathsf{T} v_\mathbf{x}$ is

$$\begin{aligned} \mathbf{u}^*(\mathbf{x}, t) &= -L(t)\,\mathbf{x} \\ L(t) &\triangleq R^{-1}B^\mathsf{T} V(t) \end{aligned}$$

The Hessian $V(t)$ and the matrix of feedback gains $L(t)$ are independent of the noise amplitude C. Thus the optimal control law $\mathbf{u}^*(\mathbf{x}, t)$ is the same for stochastic and deterministic systems (the latter is called LQR).

Linear feedback control law

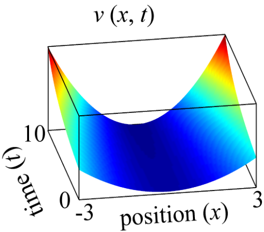
When $v(\mathbf{x}, t) = \frac{1}{2}\mathbf{x}^\mathsf{T} V(t)\,\mathbf{x} + \alpha(t)$, the optimal control $\mathbf{u}^* = -R^{-1}B^\mathsf{T} v_\mathbf{x}$ is

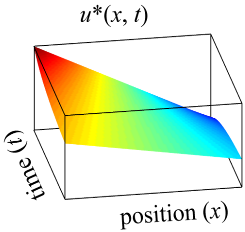
$$\begin{aligned} \mathbf{u}^*(\mathbf{x}, t) &= -L(t)\,\mathbf{x} \\ L(t) &\triangleq R^{-1}B^\mathsf{T} V(t) \end{aligned}$$

The Hessian $V(t)$ and the matrix of feedback gains $L(t)$ are independent of the noise amplitude C. Thus the optimal control law $\mathbf{u}^*(\mathbf{x}, t)$ is the same for stochastic and deterministic systems (the latter is called LQR).

Example:

$$\begin{aligned} dx &= u\,dt + 0.2\,d\omega \\ \ell(x, u) &= 0.5u^2 \\ q_T(x) &= 2.5x^2 \end{aligned}$$





LQG in discrete time

Consider an optimal control problem with dynamics and cost

$$\begin{aligned}\mathbf{x}_{k+1} &= A\mathbf{x}_k + B\mathbf{u}_k \\ \ell(\mathbf{x},\mathbf{u}) &= \frac{1}{2}\mathbf{x}^\mathsf{T} Q\mathbf{x} + \frac{1}{2}\mathbf{u}^\mathsf{T} R\mathbf{u}\end{aligned}$$

Substituting in the Bellman equation $v_k(\mathbf{x}) = \min_\mathbf{u} \{\ell(\mathbf{x},\mathbf{u}) + v_{k+1}(\mathbf{x}')\}$ and making the ansatz $v_k(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T} V_k \mathbf{x}$ yields

$$\frac{1}{2}\mathbf{x}^\mathsf{T} V_k \mathbf{x} = \min_\mathbf{u} \left\{ \frac{1}{2}\mathbf{x}^\mathsf{T} Q\mathbf{x} + \frac{1}{2}\mathbf{u}^\mathsf{T} R\mathbf{u} + \frac{1}{2}(A\mathbf{x} + B\mathbf{u})^\mathsf{T} V_{k+1}(A\mathbf{x} + B\mathbf{u}) \right\}$$

The minimum is $\mathbf{u}_k^*(\mathbf{x}) = -L_k\mathbf{x}$ where $L_k \triangleq \left(R + B^\mathsf{T} V_{k+1} B\right)^{-1} B^\mathsf{T} V_{k+1} A$.

Theorem (Riccati equation)

$$V_k = Q + A^\mathsf{T} V_{k+1}(A - BL_k)$$

Summary of Riccati equations

- Finite horizon
  - Continuous time

$$-\dot{V} = Q + A^\mathsf{T} V + VA - VBR^{-1}B^\mathsf{T} V$$

  - Discrete time

$$V_k = Q + A^\mathsf{T} V_{k+1} A - A^\mathsf{T} V_{k+1} B \left(R + B^\mathsf{T} V_{k+1} B \right)^{-1} B^\mathsf{T} V_{k+1} A$$

- Average cost
  - Continuous time ('care' in Matlab)

$$0 = Q + A^\mathsf{T} V + VA - VBR^{-1}B^\mathsf{T} V$$

  - Discrete time ('dare' in Matlab)

$$V = Q + A^\mathsf{T} VA - A^\mathsf{T} VB \left(R + B^\mathsf{T} VB \right)^{-1} B^\mathsf{T} VA$$

- Discounted cost is similar; first exit does not yield Riccati equations.

Encoding targets as quadratic costs

The matrices A, B, Q, R can be time-varying, which is useful for specifying reference trajectories $\mathbf{x}_k^*$, and for approximating non-LQG problems.

The cost $\left\|\mathbf{x}_k - \mathbf{x}_k^*\right\|^2$ can be represented in the LQG framework by augmenting the state vector as

$$\widetilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}, \quad \widetilde{A} = \begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{etc.}$$

and writing the state cost as

$$\frac{1}{2}\widetilde{\mathbf{x}}^\mathsf{T} \widetilde{Q}_k \widetilde{\mathbf{x}} = \frac{1}{2}\widetilde{\mathbf{x}}^\mathsf{T} \left(D_k^T D_k\right) \widetilde{\mathbf{x}}$$

where $D_k = \left[I, \; -\mathbf{x}_k^*\right]$ and so $D_k \widetilde{\mathbf{x}}_k = \mathbf{x}_k - \mathbf{x}_k^*$.

If the target $\mathbf{x}^*$ is stationary we can instead include it in the state, and use $D = [I, \; -I]$. This has the advantage that the resulting control law is independent of $\mathbf{x}^*$ and therefore can be used for all targets.