Linear-Quadratic-Gaussian (LQG) Controllers and Kalman Filters

Emo Todorov

Applied Mathematics and Computer Science & Engineering

University of Washington

Spring 2010

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

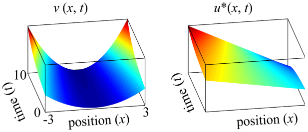
When $v(\mathbf{x},t) = \frac{1}{2}\mathbf{x}^\mathsf{T} V(t)\,\mathbf{x} + \alpha(t)$, the optimal control $\mathbf{u}^* = -R^{-1}B^\mathsf{T} v_\mathbf{x}$ is

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

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).

Example:

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



Maximum principle for LQG systems

For systems in the form

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

the optimal trajectory satisfies

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{a}(\mathbf{x}) - BR^{-1}B^\mathsf{T}\boldsymbol{\lambda} \\ -\dot{\boldsymbol{\lambda}} &= q_{\mathbf{x}}(\mathbf{x}) + \mathbf{a}_{\mathbf{x}}(\mathbf{x})^\mathsf{T}\boldsymbol{\lambda} \end{aligned}$$

Substituting $\mathbf{a}(\mathbf{x}) = A\mathbf{x}$ and $q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathsf{T} Q\mathbf{x}$ yields

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} - BR^{-1}B^\mathsf{T}\boldsymbol{\lambda} \\ -\dot{\boldsymbol{\lambda}} &= Q\mathbf{x} + A^\mathsf{T}\boldsymbol{\lambda} \end{aligned}$$

$\mathbf{x}(0)$ is given, $\boldsymbol{\lambda}(T) = Q_{\mathcal{T}}\mathbf{x}(T)$.

We can write this linear ODE as

$$\dot{\mathbf{y}} = M\mathbf{y}$$

where $\mathbf{y} = [\mathbf{x}; \boldsymbol{\lambda}]$ and

$$M = \begin{bmatrix} A & -BR^{-1}B^\mathsf{T} \\ -Q & -A^\mathsf{T} \end{bmatrix}$$

The solution is

$$\mathbf{y}(t) = \exp(Mt)\,\mathbf{y}(0)$$

Now we can solve

$$\begin{bmatrix} \mathbf{x}(T) \\ Q_{\mathcal{T}}\mathbf{x}(T) \end{bmatrix} = \exp(MT) \begin{bmatrix} \mathbf{x}(0) \\ \boldsymbol{\lambda}(0) \end{bmatrix}$$

for $\boldsymbol{\lambda}(0)$ and $\mathbf{x}(T)$.

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.

Relation between continuous and discrete time

The continuous-time system

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

can be represented in discrete time with time-step Δ as

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

In the limit $\Delta \to 0$ the discrete Riccati equation reduces to the continuous one:

$$\begin{aligned} V &= \Delta Q + (I + \Delta A)^\mathsf{T} V (I + \Delta A) \\ &\quad - (I + \Delta A)^\mathsf{T} V \Delta B \left(\Delta R + \Delta B^\mathsf{T} V \Delta B \right)^{-1} \Delta B^\mathsf{T} V (I + \Delta A) \\ V &= \Delta Q + V + \Delta A^\mathsf{T} V + \Delta VA - \Delta VB \left(R + \Delta B^\mathsf{T} VB \right)^{-1} B^\mathsf{T} V + o\left(\Delta^2\right) \\ 0 &= Q + A^\mathsf{T} V + VA - VB \left(R + \Delta B^\mathsf{T} VB \right)^{-1} B^\mathsf{T} V + \frac{1}{\Delta} o\left(\Delta^2\right) \end{aligned}$$

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.

Optimal estimation in linear-Gaussian systems

Consider the partially-observed system

$$\begin{aligned} \mathbf{x}_{k+1} &= A\mathbf{x}_k + C\boldsymbol{\omega}_k \\ \mathbf{y}_k &= H\mathbf{x}_k + D\boldsymbol{\varepsilon}_k \end{aligned}$$

with hidden state $\mathbf{x}_k$, measurement $\mathbf{y}_k$, and noise $\boldsymbol{\varepsilon}_k, \boldsymbol{\omega}_k \sim N(0, I)$.

Given a Gaussian prior $\mathbf{x}_0 \sim N(\widehat{\mathbf{x}}_0, \Sigma_0)$ and a sequence of measurements $\mathbf{y}_0, \mathbf{y}_1, \cdots \mathbf{y}_k$, we want to compute the posterior $p_{k+1}(\mathbf{x}_{k+1})$.

We can show by induction that the posterior is Gaussian at all times. Let $p_k(\mathbf{x}_k)$ be $N(\widehat{\mathbf{x}}_k, \Sigma_k)$. This will act as a prior for estimating $\mathbf{x}_{k+1}$. Now $\mathbf{x}_{k+1}$ and $\mathbf{y}_k$ are jointly Gaussian, with mean and covariance

$$\begin{aligned} E\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} &= \begin{bmatrix} A\widehat{\mathbf{x}}_k \\ H\widehat{\mathbf{x}}_k \end{bmatrix} \\ Cov\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} &= \begin{bmatrix} CC^\mathsf{T} + A\Sigma_k A^\mathsf{T} & A\Sigma_k H^\mathsf{T} \\ H\Sigma_k A^\mathsf{T} & DD^\mathsf{T} + H\Sigma_k H^\mathsf{T} \end{bmatrix} \end{aligned}$$

Kalman filter

Lemma

If $\mathbf{u}, \mathbf{v}$ *are jointly Gaussian with means* $\widehat{\mathbf{u}}, \widehat{\mathbf{v}}$ *and covariances* $\Sigma_{\mathbf{uu}}, \Sigma_{\mathbf{vv}}, \Sigma_{\mathbf{uv}} = \Sigma_{\mathbf{vu}}^\mathsf{T}$*, then* $\mathbf{u}$ *given* $\mathbf{v}$ *is Gaussian with mean and covariance*

$$
\begin{aligned}
E[\mathbf{u}|\mathbf{v}] &= \widehat{\mathbf{u}} + \Sigma_{\mathbf{uv}}\Sigma_{\mathbf{vv}}^{-1}(\mathbf{v} - \widehat{\mathbf{v}}) \\
Cov[\mathbf{u}|\mathbf{v}] &= \Sigma_{\mathbf{uu}} - \Sigma_{\mathbf{uv}}\Sigma_{\mathbf{vv}}^{-1}\Sigma_{\mathbf{vu}}
\end{aligned}
$$

Applying this to our problem with $\mathbf{u} = \mathbf{x}_{k+1}$ and $\mathbf{v} = \mathbf{y}_k$ yields

Theorem (Kalman filter)

The mean $\widehat{\mathbf{x}}$ *and covariance* Σ *of the Gaussian posterior satisfy*

$$
\begin{aligned}
\widehat{\mathbf{x}}_{k+1} &= A\widehat{\mathbf{x}}_k + K_k(\mathbf{y}_k - H\widehat{\mathbf{x}}_k) \\
\Sigma_{k+1} &= CC^\mathsf{T} + (A - K_k H)\Sigma_k A^\mathsf{T} \\
K_k &\triangleq A\Sigma_k H^\mathsf{T}\left(DD^\mathsf{T} + H\Sigma_k H^\mathsf{T}\right)^{-1}
\end{aligned}
$$

Duality of LQG control and Kalman filtering

LQG controller

State dynamics:

$$\mathbf{x}_{k+1} = (A - BL_k)\mathbf{x}_k + C\boldsymbol{\varepsilon}_k$$

Gain matrix:

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

Backward Riccati equation:

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

Kalman filter

Estimated state dynamics:

$$\widehat{\mathbf{x}}_{k+1} = (A - K_k H)\widehat{\mathbf{x}}_k + K_k\mathbf{y}_k$$

Gain matrix:

$$K_k = A\Sigma_k H^\mathsf{T}\left(DD^\mathsf{T} + H\Sigma_k H^\mathsf{T}\right)^{-1}$$

Forward Riccati equation:

$$\Sigma_{k+1} = CC^\mathsf{T} + (A - K_k H)\Sigma_k A^\mathsf{T}$$

This form of duality does not generalize to non-LQG systems. However there is a different duality which does generalize (see later). It involves an information filter, computing Σ^{-1} instead of Σ.