Control of Stochastic Processes 048913

Winter 2006 Supplement: DP for the LQ problem

20/11/2006

For the latest see http://www.ee.technion.ac.il/~adam/GRADUATES/048913

The Linear Quadratic problem is defined as follows:

1. Dynamics: $x_{t+1} = A_t x_t + B_t u_t + C_t w_t$ where u is the control, w is the noise, $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. We assume $\{w_t\}$ is i.i.d. and zero mean.
2. Cost: $V(x\pi) \stackrel{def}{=} \mathbb{E}\left[\sum_{t=0}^{N-1}(x_t^T Q_t x_t + u_t^T R_t u_t) + x_N^T Q_N x_n\right]$
   where $Q_t \geq 0$ (non-negative), $R_t > 0$ (strictly positive definite).
3. The horizon is N, initial state x_0.

With n counting "steps from 0" the DP algorithm gives

$$V_N(x) = x^T Q_N x \tag{0.1}$$

$$V_t(x) = \min_u \left[x^T Q_t x + u^T R_t u + \mathbb{E}_x^u V_{t+1}(x_1)\right] \tag{0.2}$$

$$= \min_u \left[x^T Q_t x + u^T R_t u + \mathbb{E}_x^u V_{t+1}(A_t x + B_t u + C_t w_t)\right] \tag{0.3}$$

Theorem 0.1 *V_t is quadratic in x, i.e. $V_t(x) = x^T K_t x + \alpha_t$ for some positive definite K_t and positive α_t.*

Proof. We use (backward) induction. This is trivial for $t = N$ since $V_N(x) \stackrel{def}{=} x^T Q_N x$. Assume now that $V_{t+1}(x) = x^T K_{t+1} x + \alpha_{t+1}$. By the DP equation and the induction hypothesis we have that $V_t(x)$ equals

$$
\begin{aligned}
&= \min_u \left[x^T Q_t x + u^T R_t u + \mathbb{E}_x^u \left[V_{t+1}(A_t x + B_t u + C_t w_t) \right] \right] \\
&= \min_u \left[x^T Q_t x + u^T R_t u + \mathbb{E}_x^u \left[(A_t x + B_t u + C_t w_t)^T K_{t+1} (A_t x + B_t u + C_t w_t) + \alpha + t + 1 \right] \right] \\
&= \min_u \left[x^T Q_t x + u^T R_t u + \mathbb{E}_x^u \left[(A_t x + B_t u)^T K_{t+1} (A_t x + B_t u) \right] \right] \\
&\qquad + 2(C_t w_t)^T K_{t+1} (A_t x + B_t u) + (C_t w_t)^T K_{t+1} C_t w_t + \alpha_{t+1} \big] \big] \\
&= \min_u \left[x^T Q_t x + u^T R_t u + \mathbb{E}_x^u \left[(A_t x + B_t u)^T K_{t+1} (A_t x + B_t u) \right] + 0 + \alpha_t \right]
\end{aligned}
$$

for some non negative α_t, where we obtain the last equation since $\mathbb{E}\, w_t = 0$, and since $(C_t w_t)^T K_{t+1} C_t w_t$ is positive, and therefor its expectation is non negative. So, we have

$$V_t(x) = \min_u \left[x^T Q_t x + u^T R_t u + (A_t x + B_t u)^T K_{t+1} (A_t x + B_t u) + \alpha_t \right]. \tag{0.4}$$

To minimize, differentiate with respect to u and set to 0: this gives

$$R_t u + B_t^T K_{t+1} (A_t x + B_t u) = 0 \tag{0.5}$$

and since R_t is by assumption strictly positive definite and $B_t^T K_{t+1} B_t$ is non-negative, their sum $M \stackrel{def}{=} R_t + B_t^T K_{t+1} B_t$ is invertible and we have

$$u^* = -(R_t + B_t^T K_{t+1} B_t)^{-1} B_t^T K_{t+1} A_t x \tag{0.6}$$

as the minimizer. Substituting back into the equation for $V_t(x)$ we have

$$V_t(x) = x^T K_t x + \alpha_t \tag{0.7}$$

where

$$K_t \stackrel{def}{=} Q_t + A_t^T [K_{t+1} - K_{t+1} B_t M_t^{-1} B_t^T K_{t+1}] A_t a\ . \tag{0.8}$$

The optimal decision rule at time t is thus

$$\mu_t(x) \stackrel{def}{=} L_t x = -M_t^{-1} B_t^T K_{t+1} A_t x \tag{0.9}$$

and where K_t can be computed recursively via (0.8)—the discrete-time "Riccati equation".