EE365: Linear Exponential Quadratic Regulator

Linear exponential quadratic regulator

Solution via dynamic programming

Example

Derivation of DP for LEQR

Linear exponential quadratic regulator

Linear dynamics, quadratic costs

- linear dynamics: $x_{t+1} = A_t x_t + B_t u_t + w_t$
  - $w_t \sim \mathcal{N}(0, W_t)$, $x_0 \sim \mathcal{N}(0, X_0)$ (yes, they need to be Gaussian)
  - $x_0, w_0, \ldots, w_{T-1}$ independent
- stage cost is (convex quadratic)

$$g_t(x,u) = (1/2)(x^T Q_t x + u^T R_t u)$$

  with $Q_t \geq 0$, $R_t > 0$
- terminal cost $g_T(x) = (1/2)x^T Q_T x$, $Q_T \geq 0$
- cost $C = \sum_{t=0}^{T-1} g_t(x_t, u_t) + g_T(x_T)$
- state feedback: $u_t = \mu_t(x_t)$, $t = 0, \ldots, T-1$

Linear exponential quadratic regulator

- exponential risk aversion cost

$$J = \frac{1}{\gamma} \log \mathbf{E} \exp \gamma C = R_\gamma(C)$$

  with $\gamma > 0$

- LEQR problem: choose policy $\mu_0, \ldots, \mu_{T-1}$ to minimize J
- reduces to LQR problem as $\gamma \to 0$
- for γ too large, $J = \infty$ for all policies ('neurotic breakdown')

Solution via dynamic programming

Generic risk averse dynamic programming

- optimal policy $\mu^\star$ is

$$\mu_t^\star(x) \in \underset{u}{\operatorname{argmin}} \left(g_t(x,u) + R_\gamma V_{t+1}(f_t(x,u,w_t))\right)$$

  where expectation in R_γ is over w_t

- (backward) recursion for V_t:

$$V_t(x) = \min_u \left(g_t(x,u) + R_\gamma V_{t+1}(f_t(x,u,w_t))\right)$$

DP for LEQR

we will see that

- V_t are convex quadratic (with no linear term):

$$V_t(x) = (1/2)(x^T P_t x + r_t)$$

  with $P_t \geq 0$

- optimal policy is linear: $\mu_t^\star(x) = K_t x$ (so x_t, u_t are Gaussian)
- $J = \infty$ for $\gamma \geq \gamma^{\text{crit}}$ ('neurotic breakdown')

Modified Riccati recursion

- modified Riccati recursion:

$$\begin{aligned}\tilde{P}_{t+1} &= P_{t+1} + \gamma P_{t+1}(W_t^{-1} - \gamma P_{t+1})^{-1} P_{t+1} \\ K_t &= -(R_t + B_t^T \tilde{P}_{t+1} B_t)^{-1} B_t^T \tilde{P}_{t+1} A_t \\ r_t &= r_{t+1} - (1/\gamma) \log\det(I - \gamma P_{t+1} W_t) \\ P_t &= Q_t + K_t^T R_t K^t + (A_t + B_t K_t)^T \tilde{P}_{t+1} (A_t + B_t K_t)\end{aligned}$$

- neurotic breakdown occurs if $W_t^{-1} - \gamma P_{t+1} \not> 0$ for any t
- as $\gamma \to 0$, $\tilde{P}_{t+1} \to P_{t+1}$ and

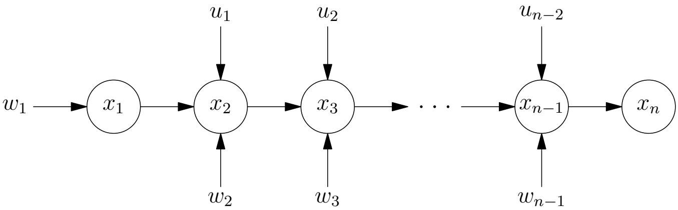
$$-(1/\gamma) \log\det(I - \gamma P_{t+1} W_t) \to \mathbf{Tr}\, P_{t+1} W_t$$

  and we recover the standard (LQR) Riccati recursion
- as in LQR, r_t keeps track of cost, doesn't affect policy

Example

LEQR example

- dynamics and actuators:



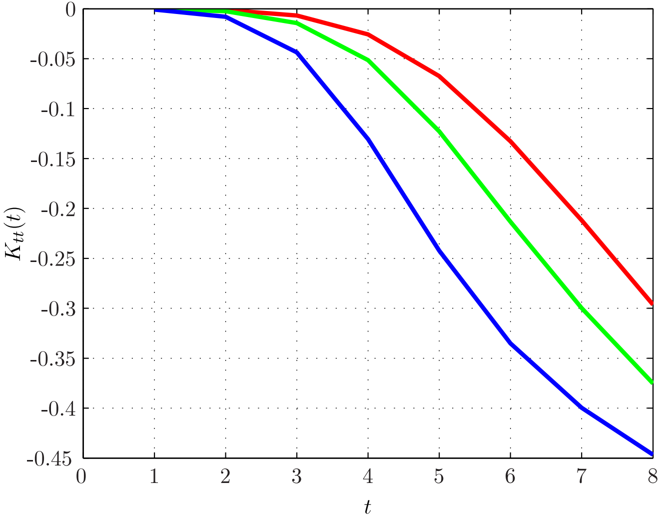
- $Q_0 = \cdots = Q_{T-1} = 0$, $Q_T = e_n e_n^T$
- $R_0 = \cdots = R_{T-1}$ diagonal with increasing values on diagonal
- $W_0 = \cdots = W_{T-1}$ diagonal with decreasing values on diagonal
- $X_0 = I$

LEQR example

γ	$\mathbf{E}\,C$	$\mathbf{std}\,C$	$R_0(C)$	$R_{1.25}(C)$	$R_{2.00}(C)$
0.00	0.3202	0.3866	**0.3202**	0.5106	1.0616
1.25	0.3347	0.3543	0.3347	**0.4740**	0.7918
2.00	0.3858	0.3475	0.3885	0.5099	**0.7353**

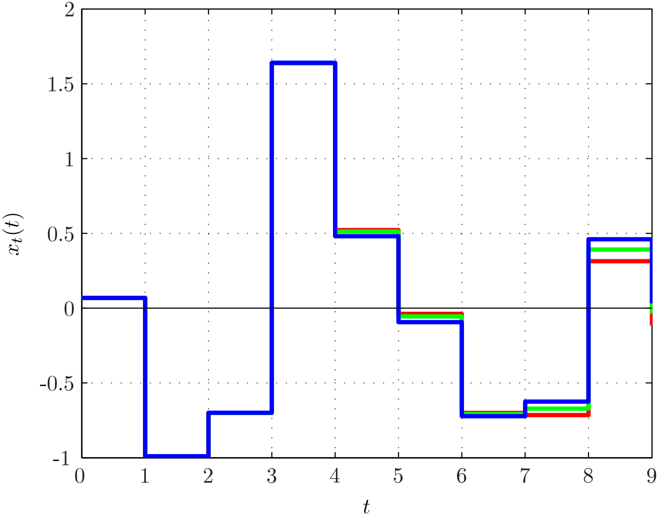
LEQR example

feedback gain: $\gamma = 0$, $\gamma = 1.2$, $\gamma = 2.0$



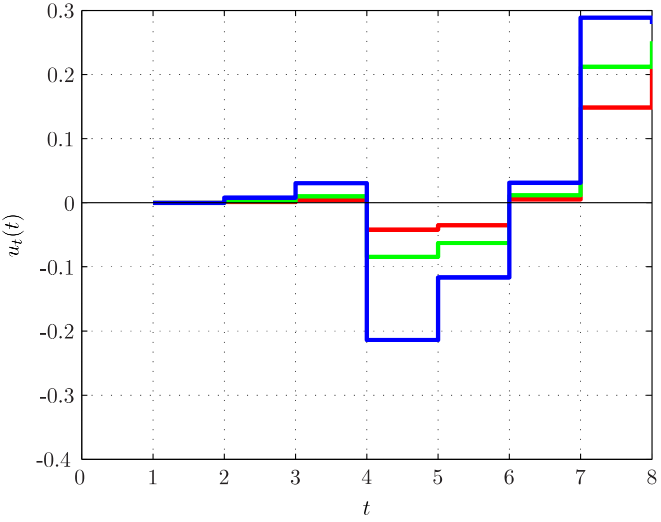
LEQR example

sample realization (state): $\gamma = 0$, $\gamma = 1.2$, $\gamma = 2.0$



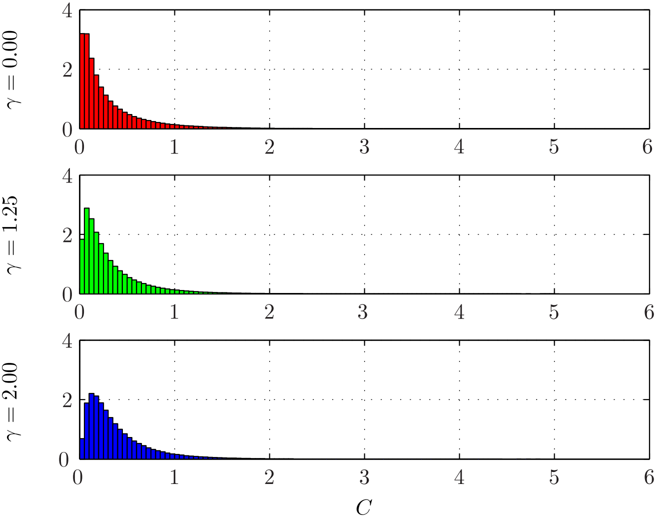
LEQR example

sample realization (input): $\gamma = 0$, $\gamma = 1.2$, $\gamma = 2.0$



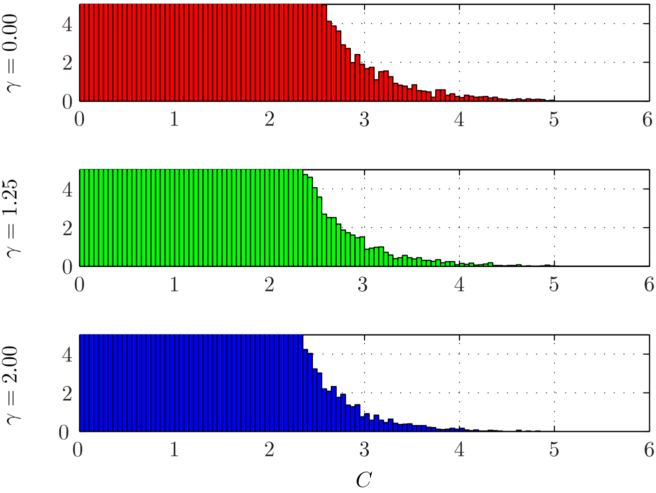
LEQR example

cost histogram



LEQR example

cost histogram (tails)



Derivation of DP for LEQR

Expectation of exponential of quadratic of Gaussian

- suppose $z \sim \mathcal{N}(\bar{z}, Z)$, $P > 0$
- $\mathbf{E}(1/2)z^T Pz = (1/2)\,\mathbf{Tr}\,PZ$
- let $J = R_\gamma(z^T Pz/2) = \frac{1}{\gamma}\log \mathbf{E}\exp(\gamma/2)z^T Pz$
- then $J = \infty$ if $Z^{-1} \not> \gamma P$
- when $Z^{-1} > \gamma P$,

$$J = \frac{1}{2}\left(\bar{z}^T \tilde{P}\bar{z} - (1/\gamma)\log\det(I - \gamma PZ)\right)$$

  where $\tilde{P} = P + \gamma P(Z^{-1} - \gamma P)^{-1} P$
- as $\gamma \to 0$, $\tilde{P} \to P$, $J \to (1/2)\,\mathbf{Tr}\,PZ$

Derivation

- to get formula above start with integral

$$\mathbf{E}\exp(\gamma/2)z^TPz = \frac{1}{(2\pi)^{n/2}(\det Z)^{1/2}}\int e^{\gamma x^TPx/2}\; e^{-(x-\bar{z})^TZ^{-1}(x-\bar{z})/2}\; dx$$

- simplify integrand, complete squares, and use

$$\frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}}\int e^{-(x-\mu)^T\Sigma^{-1}(x-\mu)/2}\; dx = 1$$

  to get formula above

Limit

For any $Z \in \mathbb{R}^{n\times n}$

$$\lim_{\gamma\to 0} -\frac{1}{\gamma}\log\det(I-\gamma Z) = \mathbf{Tr}(Z).$$

let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of Z, then

$$\begin{aligned} -\frac{1}{t}\log\det(I-tZ) &= -\frac{1}{t}\log\prod_{i=1}^{n}(1-t\lambda_i) = -\frac{1}{t}\sum_{i=1}^{n}\log(1-t\lambda_i) \\ &= \frac{1}{t}\sum_{i=1}^{n}\sum_{k=1}^{\infty}\frac{1}{k}(t\lambda_i)^k \\ &= \sum_{k=1}^{\infty}\frac{1}{k}\sum_{i=1}^{n}\lambda_i^k t^{k-1} \\ &= \sum_{k=0}^{\infty}\left(\frac{1}{k+1}\sum_{i=1}^{n}\lambda_i^{k+1}\right)t^k \end{aligned}$$

as $t \to 0$, we are left with the term corresponding to $k = 0$

$$\lim_{t\to\infty} -\frac{1}{t}\log\det(I-tZ) = \sum_{i=1}^{n}\lambda_i = \mathbf{Tr}(Z).$$

Derivation of DP for LEQR

- proof by induction: suppose $V_{t+1}(x) = (1/2)(x^T P_{t+1} x + r_{t+1})$
- we need to minimize over u

$$\begin{array}{l} g_t(x,u) + R_\gamma(V_{t+1}(f_t(x,u,w_t))) \\ = (1/2)(x^T Q_t x + u^T R_t u) \\ \quad + R_\gamma\left((1/2)((A_t x + B_t u + w_t)^T P_{t+1}(A_t x + B_t u + w_t) + r_{t+1})\right) \end{array}$$

- same as minimizing

$$(1/2)u^T R_t u + R_\gamma\left((1/2) z^T P_{t+1} z\right)$$

where $z \sim \mathcal{N}(A_t x + B_t u, W_t)$

Derivation of DP for LEQR

- using formula for $R_\gamma\left((1/2)z^T P_{t+1} z\right)$ above, need to minimize over u

$$\frac{1}{2}\left(u^T R_t u + (A_t x + B_t u)^T \tilde{P}(A_t x + B_t u) - (1/\gamma)\log\det(I - \gamma Z P)\right)$$

  where $\tilde{P} = P + \gamma P(Z^{-1} - \gamma P)^{-1} P$

- this expression is ∞ if $Z^{-1} \not> \gamma P$

- otherwise: last term is constant, so

$$\mu_t^\star(x) = -(R_t + B_t^T \tilde{P}_{t+1} B_t)^{-1} B_t^T \tilde{P}_{t+1} A_t x$$

- adding back in constant terms to get V_t, we get modified Riccati recursion given above