EE365: Linear Quadratic Stochastic Control

Continuous state Markov decision process

Affine and quadratic functions

Linear quadratic Markov decision process

Linear quadratic regulator

Linear quadratic trading

Outline

Continuous state Markov decision problem

- dynamics: $x_{t+1} = f_t(x_t, u_t, w_t)$
- $x_0, w_0, w_1, \ldots$ independent
- stage cost: $g_t(x_t, u_t, w_t)$
- (state feedback) policy: $u_t = \mu_t(x_t)$
- choose policy to minimize

$$J = \mathbf{E} \left(\sum_{t=0}^{T-1} g_t(x_t, u_t, w_t) + g_T(x_T) \right)$$

- we consider the case $\mathcal{X} = \mathbf{R}^n$, $\mathcal{U} = \mathbf{R}^m$

Continuous state Markov decision problem

- many (mostly mathematical) pathologies can occur in this case
  - but not in the special case we'll consider
- a basic issue: how do you even *represent* the functions f_t, g_t, and μ_t?
  - for n and m very small (say, 2 or 3) we can use *gridding*
  - we can give the coefficients in some (dense) basis of functions
  - most generally, we assume we have a method to compute function values, given the arguments
  - exponential growth that occurs in gridding is called *curse of dimensionality*

Continuous state Markov decision problem: Dynamic programming

- set $V_T(x) = g_T(x)$
- for $t = T-1, \ldots, 0$,

$$\mu_t(x) \in \mathrm{argmin}_u\, \mathbf{E}\left(g_t(x,u,w_t) + V_{t+1}(f_t(x,u,w_t))\right)$$
$$V_t(x) = \mathbf{E}\left(g_t(x,\mu_t(x),w_t) + V_{t+1}(f_t(x,\mu_t(x),w_t))\right)$$

- this gives value functions and optimal policy, **in principle only**
- but you can't in general represent, much less compute, V_t or μ_t

Continuous state Markov decision problem: Dynamic programming

for DP to be tractable, f_t and g_t need to have special form for which we can

- represent V_t, μ_t in some tractable way
- carry out expectation and minimization in DP recursion

one of the few situations where this holds: **linear quadratic problems**

- f_t is an affine function of x_t, u_t ('linear dynamical system')
- g_t are convex quadratic functions of x_t, u_t

Linear quadratic problems

for linear quadratic problems

- value functions $V_t^\star$ are quadratic
- hence representable by their coefficients
- we can carry out the expectation and the minimization in DP recursion explicitly **using linear algebra**
- optimal policy functions are affine: $\mu_t^\star(x) = K_t x + l_t$
- we can compute the coefficients K_t and l_t explicitly

in other words:
we can solve linear quadratic stochastic control problems in practice

Outline

Affine functions

- $f : \mathbf{R}^p \to \mathbf{R}^q$ is affine if it has the form

$$f(x) = Ax + b$$

  i.e., it is a linear function plus a constant

- a linear function is special case, with $b = 0$
- affine functions closed under sum, scalar multiplication, composition (with explicit formulas for coefficients in each case)

Quadratic function

- $f : \mathbf{R}^n \to \mathbf{R}$ is quadratic if it has the form

$$f(x) = (1/2)x^T Px + q^T x + (1/2)r$$

  with $P = P^T \in \mathbf{R}^{n\times n}$ (the 1/2 on r is for convenience)

- often write as quadratic form in $(x, 1)$:

$$f(x) = (1/2) \left[\begin{array}{c} x \\ 1 \end{array}\right]^T \left[\begin{array}{cc} P & q \\ q^T & r \end{array}\right] \left[\begin{array}{c} x \\ 1 \end{array}\right]$$

- special cases:
  - quadratic form: $q = 0$, $r = 0$
  - affine (linear) function: $P = 0$ ($P = 0$, $r = 0$)
  - constant: $P = 0$, $q = 0$
- uniqueness: $f(x) = \tilde{f}(x) \iff P = \tilde{P},\ q = \tilde{q},\ r = \tilde{r}$

Calculus of quadratic functions

- quadratic functions on $\mathbf{R}^n$ form a vector space of dimension

$$\frac{n(n+1)}{2} + n + 1$$

- *i.e.*, they are closed under addition, scalar multiplication

Composition of quadratic and affine functions

- suppose
  - $f(z) = (1/2)z^T Pz + q^T z + (1/2)r$ is quadratic function on $\mathbf{R}^m$
  - $g(x) = Ax + b$ is affine function from $\mathbf{R}^n$ into $\mathbf{R}^m$
- then composition $h(x) = (f \circ g)(x) = f(Ax + b)$ is quadratic
- write $h(x)$ as

$$(1/2) \begin{bmatrix} x \\ 1 \end{bmatrix}^T \left(\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}^T \begin{bmatrix} P & q \\ q^T & r \end{bmatrix} \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ 1 \end{bmatrix}$$

- so matrix multiplication gives us the coefficient matrix of h

Convexity and nonnegativity of a quadratic function

- f is convex (graph does not curve down) if and only if $P \geq 0$ (matrix inequality)
- f is strictly convex (graph curves up) if and only if $P > 0$ (matrix inequality)
- f is nonnegative (*i.e.*, $f(x) \geq 0$ for all x) if and only if

$$\begin{bmatrix} P & q \\ q^T & r \end{bmatrix} \geq 0$$

- $f(x) > 0$ if and only if matrix inequality is strict
- nonnegative $\Rightarrow$ convex

Checking convexity and nonnegativity

- we can check convexity or nonnegativity in $O(n^3)$ operations by eigenvalue decomposition, Cholesky factorization, ...
- composition with affine function preserves convexity, nonnegativity:

$$f \text{ convex}, \ g \text{ affine} \implies f \circ g \text{ convex}$$

- linear combination of convex quadratics, with nonnegative coefficients, is convex quadratic
- if $f(x, w)$ is convex quadratic in x for each w (a random variable) then

$$g(x) = \mathop{\mathbf{E}}_{w} f(x, w)$$

  is convex quadratic (*i.e.*, convex quadratics closed under expectation)

Minimizing a quadratic

- if f is not convex, then $\min_x f(x) = -\infty$
- otherwise, x minimizes f if and only if $\nabla f(x) = Px + q = 0$
- for $q \notin \text{range}(P)$, $\min_x f(x) = -\infty$
- for $P > 0$, unique minimizer is $x = -P^{-1}q$
- minimum value is

$$\min_x f(x) = -(1/2)q^T P^{-1} q + (1/2)r$$

  (a concave quadratic function of q)
- for case $P \geq 0$, $q \in \text{range}(P)$, replace P^{-1} with $P^\dagger$

Partial minimization of a quadratic

- suppose f is a quadratic function of (x, u), convex in u
- then the partial minimization function

$$g(x) = \min_u f(x, u)$$

  is a quadratic function of x; if f is convex, so is g
- the minimizer $\mathop{\mathrm{argmin}}_u f(x, u)$ is an affine function of x
- minimizing a convex quadratic function over some variables yields a convex quadratic function of the remaining ones
- *i.e.*, convex quadratics closed under partial minimization

Partial minimization of a quadratic

- let's take

$$f(x,u) = (1/2)\begin{bmatrix} x \\ u \\ 1 \end{bmatrix}^T \begin{bmatrix} P_{xx} & P_{xu} & q_x \\ P_{ux} & P_{uu} & q_u \\ q_x^T & q_u^T & r \end{bmatrix} \begin{bmatrix} x \\ u \\ 1 \end{bmatrix}$$

  with $P_{uu} > 0$, $P_{ux} = P_{xu}^T$

- minimizer of f over u satisfies

$$0 = \nabla_u f(x,u) = P_{uu}u + P_{ux}x + q_u$$

  so $u = -P_{uu}^{-1}(P_{ux}x + q_u)$ is an affine function of x

Partial minimization of a quadratic

- substituting u into expression for f gives

$$g(x) = (1/2) \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} P_{xx} - P_{xu} P_{uu}^{-1} P_{ux} & q_x - P_{xu} P_{uu}^{-1} q_u \\ q_x^T - q_u^T P_{uu}^{-1} P_{ux} & r - q_u P_{uu}^{-1} q_u \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}$$

- $P_{xx} - P_{xu} P_{uu}^{-1} P_{ux}$ is the Schur complement of P w.r.t. u
- $P_{xx} - P_{xu} P_{uu}^{-1} P_{ux} \geq 0$ if $P \geq 0$
- or simpler: g is composition of f with affine function $x \mapsto (x, u)$

$$\begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} I \\ -P_{uu}^{-1} P_{ux} \end{bmatrix} x + \begin{bmatrix} 0 \\ -P_{uu}^{-1} q_u \end{bmatrix}$$

  - we already know how to form composition quadratic (affine)
  - and the result is convex

Summary

convex quadratics are closed under

- addition
- expectation
- pre-composition with an affine function
- partial minimization

in each case, we can explicitly compute the coefficients of the result using linear algebra

Outline

(Random) linear dynamical system

- dynamics $x_{t+1} = f_t(x_t, u_t, w_t) = A_t(w_t)x_t + B_t(w_t)u_t + c_t(w_t)$
- for each w_t, f_t is affine in (x_t, u_t)
- $x_0, w_0, w_1, \ldots$ are independent
- $A_t(w_t) \in \mathbf{R}^{n \times n}$ is dynamics matrix
- $B_t(w_t) \in \mathbf{R}^{n \times m}$ is input matrix
- $c_t(w_t) \in \mathbf{R}^n$ is offset

Linear quadratic stochastic control problem

- stage cost $g_t(x_t, u_t, w_t)$ is convex quadratic in (x_t, u_t) for each w_t
- choose policy $u_t = \mu_t(x_t)$ to minimize objective

$$J = \mathbf{E}\left(\sum_{t=0}^{T-1} g_t(x_t, u_t, w_t) + g_T(x_T)\right)$$

Dynamic programming

- set $V_T(x) = g_T(x)$
- for $t = T-1, \ldots, 0$,

$$\mu_t(x) \in \operatorname{argmin}_u \mathbf{E}\left(g_t(x,u,w_t) + V_{t+1}(f_t(x,u,w_t))\right)$$
$$V_t(x) = \mathbf{E}\left(g_t(x,\mu_t(x),w_t) + V_{t+1}(f_t(x,\mu_t(x),w_t))\right)$$

- all V_t are convex quadratic, and all μ_t are affine
- this gives value functions and optimal policy, **explicitly**

Dynamic programming

we show V_t are convex quadratic by (backward) induction

- suppose $V_T, \ldots, V_{t+1}$ are convex quadratic
- since f_t is affine in (x, u), $V_{t+1}(f_t(x, u, w_t))$ is convex quadratic
- so $g_t(x, u, w_t) + V_{t+1}(f_t(x, u, w_t))$ is convex quadratic
- and so is its expectation over w_t
- partial minimization over u leaves convex quadratic of x, which is $V_t(x)$
- argmin is affine function of x, so optimal policy is affine

Linear equality constraints

- can add (deterministic) linear equality constraints on x_t, u_t into $g_t,\ g_T$:

$$g_t(x, u, w) = g_t^{\mathrm{quad}}(x, u, w) + \begin{cases} 0 & F_t x + G_t u = h_t \\ \infty & \text{otherwise} \end{cases}$$

- everything still works:
  - V_t is convex quadratic, possibly with equality constraints
  - μ_t is affine
- reason: minimizing a convex quadratic over some variables, subject to equality constraints, yields a convex quadratic in remaining variables

Infinite horizon linear quadratic problems

- consider average stage cost problems (others are similar) with time-invariant dynamics and stage costs
- same as for finite state case: use value iteration
- set $V_0(x) = 0$; for $k = 0, 1, \ldots$,

$$\mu_{k+1}(x) = \mathop{\mathrm{argmin}}_u \mathbf{E}\left(g(x, u, w_t) + V_k(f(x, u, w_t))\right)$$
$$V_{k+1}(x) = \mathbf{E}\left(g(x, \mu_{k+1}(x), w_t) + V_k(f(x, \mu_{k+1}(x), w_t))\right)$$

- can be carried out concretely, since V_k is quadratic, μ_k is affine

Optimal steady-state policy

- $\mu_k \to \mu^\star$ (ITAP), a.k.a. steady-state policy $\mu^\star(x) = K^\star x + l^\star$
- $K^\star$ ($l^\star$) called (steady-state, average cost) optimal gain matrix (offset)
- $V_k(x) - V_k(x') \to V^{\mathrm{rel}}(x)$, relative value function (ITAP)
  - x' is (arbitrary) reference state
  - V^{rel} defined only up to a constant
- $V_{k+1}(x) - V_k(x) \to J^\star$, the optimal average cost, for any x

Outline

Linear quadratic regulator

- $x_{t+1} = A_t x_t + B_t u_t + w_t$
- $\mathbf{E}\, w_t = 0$, $\mathbf{E}\, w_t w_t^T = W_t$
- stage cost is (convex quadratic)

  $$(1/2)(x_t^T Q_t x_t + u_t^T R_t u_t)$$

  with $Q_t \geq 0$, $R_t > 0$
- terminal cost $(1/2) x_T^T Q_T x_T$, $Q_T \geq 0$
- variation: terminal constraint $x_T = 0$

Linear quadratic regulator: DP

- value functions are quadratic plus constant (linear terms are zero):

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

- $P_T = Q_T$, $r_T = 0$
- optimal expected tail cost:

$$\begin{aligned}
&\mathbf{E}\, V_{t+1}(f_t(x, u, w_t)) \\
&\quad = (1/2)(r_{t+1} + \mathbf{E}(A_t x + B_t u + w_t)^T P_{t+1}(A_t x + B_t u + w_t)) \\
&\quad = (1/2)(r_{t+1} + (A_t x + B_t u)^T P_{t+1}(A_t x + B_t u) + \mathbf{Tr}(P_{t+1} W_t))
\end{aligned}$$

using $\mathbf{E}\, w_t = 0$ and

$$\mathbf{E}\, w_t^T P_{t+1} w_t = \mathbf{E}\,\mathbf{Tr}(P_{t+1} w_t w_t^T) = \mathbf{Tr}(P_{t+1} W_t)$$

Linear quadratic regulator: DP

- minimize over u to get optimal policy:

$$\begin{aligned}\mu_t(x) &= \operatorname*{argmin}_u \left(u^T R_t u + u^T B_t^T P_{t+1} B_t u + 2(B_t^T P_{t+1} A_t x)^T u\right) \\ &= -\left(R_t + B_t^T P_{t+1} B_t\right)^{-1} B_t^T P_{t+1} A_t x \\ &= K_t x\end{aligned}$$

- optimal policy is linear (as opposed to affine)
- using $u = K_t x$ we then have

$$V_t(x) = (1/2)(r_{t+1} + \mathbf{Tr}(P_{t+1} W_t) + x^T(Q_t + K_t^T R_t K_t)x + x^T(A_t + B_t K_t)^T P_{t+1}(A_t + B_t K_t)x)$$

- so coefficients of V_t are

$$\begin{aligned}P_t &= Q_t + K_t^T R_t K_t + (A_t + B_t K_t)^T P_{t+1}(A_t + B_t K_t), \\ r_t &= r_{t+1} + \mathbf{Tr}(P_{t+1} W_t)\end{aligned}$$

Linear quadratic regulator: Riccati recursion

- set $P_T = Q_T$
- for $t = T-1, \ldots, 0$

$$\begin{aligned} K_t &= -(R_t + B_t^T P_{t+1} B_t)^{-1} B_t^T P_{t+1} A_t \\ P_t &= Q_t + K_t^T R_t K_t + (A_t + B_t K_t)^T P_{t+1} (A_t + B_t K_t) \end{aligned}$$

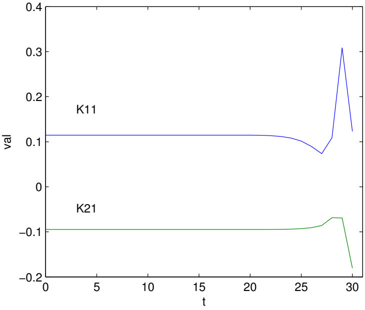
- called Riccati recursion; gives optimal policies, which are linear functions
- **surprise:** optimal policy does not depend on the disturbance distribution (provided it is zero mean)
- $J^\star = (1/2)(\mathbf{Tr}(P_0 X_0) + \sum_{t=0}^{T-1} \mathbf{Tr}(P_{t+1} W_t))$, where $X_0 = \mathbf{E}(x_0 x_0^T)$

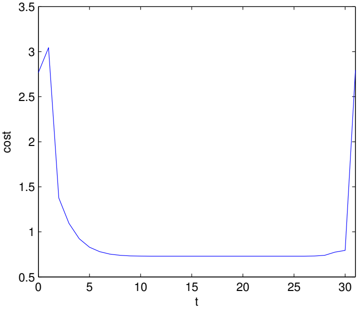
Linear quadratic regulator: Example

- $n = 5$ states, $m = 2$ inputs, horizon $T = 31$
- A, B chosen randomly; A scaled so $\max_i |\lambda_i(A)| = 1$
- $Q_t = I$, $R_t = I$, $t = 0, \ldots, T-1$, $Q_T = 5I$
- $x_0 \sim \mathcal{N}(0, X_0)$, $X_0 = I$
- $w_t \sim \mathcal{N}(0, W)$, $W = 0.1I$

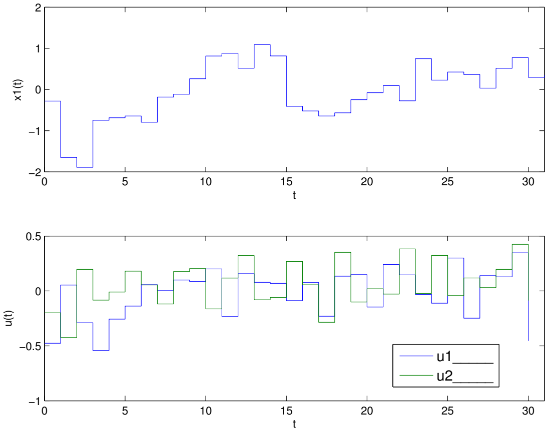
Linear quadratic regulator: Example

left: $(K_t)_{11}$, $(K_t)_{21}$ vs. t; *right*: $\mathbf{E}\, J_t$ vs. t





Linear quadratic regulator: Sample trajectory



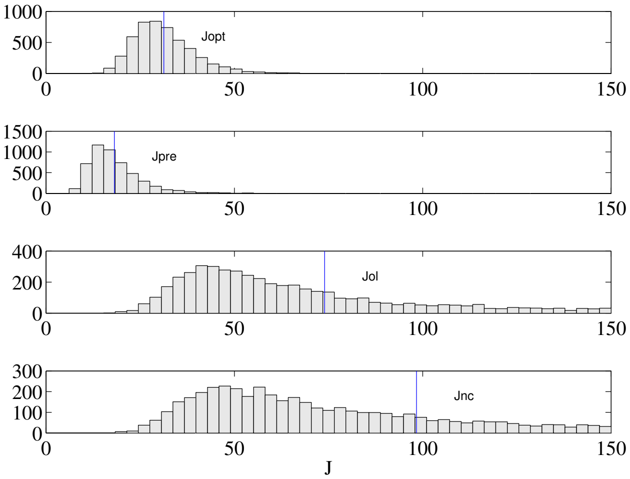
Linear quadratic regulator: Cost comparison

compare cost for

- optimal policy, $J^\star$
- prescient policy, J^{pre}: $w_0 \ldots, w_T$ known in advance
- open loop policy, J^{ol}: choose $u_0, \ldots, u_T$ with knowledge of x_0 only
- no control (1-step greedy), J^{nc}: $u_0, \ldots, u_T = 0$

Linear quadratic regulator: Cost comparison

total stage cost histograms, $N = 5000$ Monte Carlo simulations



Steady-state linear quadratic regulator

- average cost case, all data time-invariant
- use Riccati recursion to find steady-state (average cost) optimal policy:

$$
\begin{aligned}
K_{k+1} &= -(R + B^T P_k B)^{-1} B^T P_k A \\
P_{k+1} &= Q + K_{k+1}^T R K_{k+1} + (A + BK_{k+1})^T P_k (A + BK_{k+1})
\end{aligned}
$$

- $K_k \to K^\star$, steady-state (average cost) optimal gain: $\mu^\star(x) = K^\star x$
- $(1/2)x^T P_k x \to V^{\mathrm{rel}}(x)$ with reference state $x' = 0$
- $(1/2)\,\mathbf{Tr}(P_k W) \to J^\star$, optimal average stage cost

Outline

Continuous state Markov decision process

Affine and quadratic functions

Linear quadratic Markov decision process

Linear quadratic regulator

Linear quadratic trading

Linear quadratic trading: Dynamics

- $x_{t+1} = f_t(x_t, u_t, \rho_t) = \mathbf{diag}(\rho_t)(x_t + u_t)$
- $x_t \in \mathbf{R}^n$ is dollar amount of holding in n assets
- $(x_t)_i < 0$ means short position in asset i in period t
- $u_t \in \mathbf{R}^n$ is dollar amount of each asset bought at beginning of period t
- $(u_t)_i < 0$ means asset i is sold in period t
- $x_t^+ = x_t + u_t$ is post-trade portfolio
- $\rho_t \in \mathbf{R}^n_{++}$ is (random) return of assets over period $(t, t+1]$
- returns independent, with $\mathbf{E}\,\rho_t = \overline{\rho}_t$, $\mathbf{E}\,\rho_t\rho_t^T = \Sigma_t$

Linear quadratic trading: Stage cost

stage cost for $t = 0, \ldots, T-1$ is (convex quadratic)

$$g_t(x, u) = \mathbf{1}^T u + (1/2)(\kappa_t^T u^2 + \gamma(x+u)^T Q_t (x+u))$$

with $Q_t > 0$

- first term is gross cash in
- second term is quadratic transaction cost (square is elementwise; $\kappa_t > 0$)
- third term is risk (variance of post-trade portfolio for $Q_t = \Sigma_t - \overline{\rho}_t \overline{\rho}_t^T$)
- $\gamma > 0$ is risk aversion parameter
- minimizing total stage cost equivalent to maximizing (risk-penalized) net cash taken from portfolio

Linear quadratic trading: Terminal cost

- terminal cost: $g_T(x) = -\mathbf{1}^T x + (1/2)\kappa_T^T x^2$, $\kappa_T > 0$
- this is net cash in if we close out (liquidate) final positions, with quadratic transaction cost

Linear quadratic trading: DP

- value functions quadratic (including linear and constant terms):

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

- we'll need formula

$$\mathbf{E}(\mathbf{diag}(\rho_t) P \, \mathbf{diag}(\rho_t)) = P \circ \Sigma_t$$

  where $\circ$ is Hadamard (element-wise) product

- optimal expected tail cost

$$\begin{aligned}\mathbf{E}\, V_{t+1}(f_t(x,u,\rho_t)) &= \mathbf{E}\, V_{t+1}(\mathbf{diag}(\rho_t)x^+) \\ &= (1/2)((x^+)^T P_{t+1} \circ \Sigma_t x^+ + 2q_{t+1}^T \mathbf{diag}(\bar{\rho}_t)x^+ + r_{t+1})\end{aligned}$$

Linear quadratic trading: DP

- $P_T = \mathbf{diag}(\kappa_T)$, $q_T = -\mathbf{1}$, $r_T = 0$
- recall $V_t(x) = \min_u \mathbf{E}\left(g_t(x,u) + V_{t+1}(\mathbf{diag}(\rho_t)(x+u))\right)$
- for $t = T-1, \ldots, 0$ we minimize over u to get optimal policy:

$$\begin{aligned} \mu_t(x) &= \mathop{\mathrm{argmin}}_u \left(u^T(S_{t+1} + \mathbf{diag}(\kappa_t))u + 2(S_{t+1}x + s_{t+1} + \mathbf{1})^T u\right) \\ &= -(S_{t+1} + \mathbf{diag}(\kappa_t))^{-1}(S_{t+1}x + s_{t+1} + \mathbf{1}) \\ &= K_t x + l_t \end{aligned}$$

  where

$$S_{t+1} = P_{t+1} \circ \Sigma_t + \gamma Q_t, \qquad s_{t+1} = \overline{\rho}_t \circ q_{t+1}$$

- using $u = K_t x + l_t$ we then have

$$V_t(x) = (1/2) \left[\begin{array}{c} x \\ 1 \end{array}\right]^T \left[\begin{array}{cc} S_{t+1}(I + K_t) & s_{t+1} + S_{t+1} l_t \\ s_{t+1}^T + l_t^T S_{t+1} & r_{t+1} + (s_{t+1} + \mathbf{1})^T l_t \end{array}\right] \left[\begin{array}{c} x \\ 1 \end{array}\right]$$

Linear quadratic trading: Value iteration

- set $P_T = \mathbf{diag}(\kappa_T)$, $q_T = -\mathbf{1}$, $r_T = 0$
- for $t = T-1, \ldots, 0$

$$\begin{array}{rcl} K_t &=& -(S_{t+1} + \mathbf{diag}(\kappa_t))^{-1} S_{t+1} \\ l_t &=& -(S_{t+1} + \mathbf{diag}(\kappa_t))^{-1} (s_{t+1} + \mathbf{1}) \\ P_t &=& S_{t+1}(I + K_t) \\ q_t &=& s_{t+1} + S_{t+1} l_t \\ r_t &=& r_{t+1} + (s_{t+1} + \mathbf{1})^T l_t \end{array}$$

  where

$$S_{t+1} = P_{t+1} \circ \Sigma_t + \gamma Q_t, \qquad s_{t+1} = \bar{\rho}_t \circ q_{t+1}$$

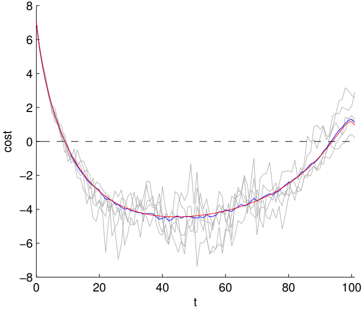
- optimal policy: $\mu_t^\star(x) = K_t x + l_t$
- can write as $\mu_t^\star(x) = K_t(x - x_t^{\mathrm{tar}})$, $x_t^{\mathrm{tar}} = -K_t^{-1} l_t = -S_{t+1}^{-1}(s_{t+1} + \mathbf{1})$
- $J^\star = \mathbf{E}\, V_0(x_0)$

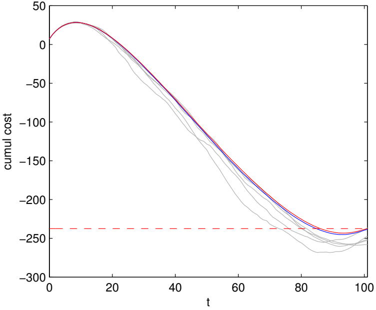
Linear quadratic trading: Numerical instance

- $n = 30$ assets over $T = 100$ time-steps
- initial portfolio $x_0 = 0$
- $\overline{\rho}_t = \overline{\rho}$, $\Sigma_t = \Sigma$ for $t = 0, \ldots, T-1$
- $Q_t = \Sigma - \overline{\rho}\overline{\rho}^T$ for $t = 0, \ldots, T-1$
- asset returns log-normal, expected returns range over $\pm 3\%$ per period
- asset return standard deviations range from 0.4% to 9.8%
- asset correlations range from -0.3 to 0.8

Linear quadratic trading: Numerical instance

- ran $N = 100$ Monte Carlo simulations
- $J^\star = V_0(x_0) = -237.5$ (Monte Carlo estimate: -238.4)
- *left*: stage cost; *right*: cumulative stage cost
- exact (red), MC estimate (blue), and samples (gray); $J^\star$ red dashed





Linear quadratic trading: Numerical instance

we define $x_{T+1} = 0$, *i.e.*, we close out the position during period T