PS 1 Problem 1

Thursday, April 4, 2019 4:35 PM

$$f_1(x,y) = -\log(10-2x^2-y^2)$$
$$f_2(x,y) = x^2(1+2y-x^2)$$

a) NOC $\Rightarrow \nabla f(x^*) = 0$

$x^* = (0,0) \qquad x=0,\ y=0$

$$\nabla f_1(0,0) = \begin{bmatrix} \frac{\partial f_1}{\partial x} \\ \frac{\partial f_1}{\partial y} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-1(-4x)}{10-2x^2-y^2} \\ \frac{-1(-2y)}{10-2x^2-y^2} \end{bmatrix} = \begin{bmatrix} \frac{4x}{10-2x^2-y^2} \\ \frac{2y}{10-2x^2-y^2} \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla^2 f_1(0,0) = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x^2} & \frac{\partial^2 f_1}{\partial x \partial y} \\ \frac{\partial^2 f_1}{\partial x \partial y} & \frac{\partial^2 f_1}{\partial y^2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{4(10-2x^2-y^2)-(4x)(-4x)}{(10-2x^2-y^2)^2} & \frac{0-(4x)(-2y)}{(10-2x^2-y^2)^2} \\ \frac{0-(2y)(-4x)}{(10-2x^2-y^2)^2} & \frac{2(10-2x^2-y^2)-(2y)(-4y)}{(10-2x^2-y^2)^2} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{4(10)}{10^2} & 0 \\ 0 & \frac{2(10)}{10^2} \end{bmatrix}$$

$$= \begin{bmatrix} 0.4 & 0 \\ 0 & 0.2 \end{bmatrix}$$

$\therefore$ eig $= 0.4,\ 0.2 \quad > 0$

$\therefore$ $\nabla^2 f_1(0,0)$ is pos def

$\Rightarrow$ also pos semi-def

f_1: $(0,0)$ satisfies NOC to 2^{nd} order ($\nabla f_1 = 0$ $\nabla^2 f_1$ pos semi-def) for local min

$(0,0)$ also satisfies SOC ($\nabla f_1 = 0$ $\nabla^2 f_1$ pos def) for local min

$$\nabla f_2(0,0) = \begin{bmatrix} \frac{\partial f_2}{\partial x} \\ \frac{\partial f_2}{\partial y} \end{bmatrix}$$

$$= \begin{bmatrix} 2x + 4xy - 4x^3 \\ 2x^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla^2 f_2(0,0) = \begin{bmatrix} \frac{\partial^2 f_2}{\partial x^2} & \frac{\partial^2 f_2}{\partial x \partial y} \\ \frac{\partial^2 f_2}{\partial x \partial y} & \frac{\partial^2 f_2}{\partial y^2} \end{bmatrix}$$

$$= \begin{bmatrix} 2 + 4y - 12x^2 & 4x \\ 4x & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

eig are 2 & 0

$\Rightarrow$ $\nabla^2 f_2$ is pos semi-def

NOT pos def

$\boxed{\begin{array}{l} f_2 : (0,0) \text{ satisfies NOC to } 2^{nd} \text{ order} \\ \quad (\nabla f_2 = 0 \quad \nabla^2 f_2 \text{ pos semi-def}) \text{ for local min} \\ \quad (0,0) \text{ does NOT satisfy SOC} \\ \quad (\nabla f_1 = 0 \quad \nabla^2 f_1 \text{ not pos def}) \text{ for local min} \end{array}}$

b) f_1 : $(0,0)$ is a local $\min^m$ $\because$ SOC satisfied

$(0,0)$ is also global $\min^m$ $\because f_1$ convex

f_2 : $(0,0)$ is NOT a local $\min^m$

@ $(0,0)$ $f_2 = 0$

@ $(10,0)$ $f_2 = 100\,(1 + 0 - 100) = -9900$

@ $(\varepsilon, -\gamma)$ $f_2 = \varepsilon^2 (1 - \gamma - \varepsilon^2)$

$f_2(\varepsilon, -\gamma) < f_2(\varepsilon, 0)$

$\therefore$ NOT a local $\min^m$

$\therefore$ Also not a global $\min^m$

PS 1 Problem 2

Thursday, April 4, 2019 5:05 PM

$\min \quad x_1 + x_2$

$\text{st} \quad x_1^2 + x_2^2 = 2$

$\mathcal{L} = x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 2)$

Let x^*, λ^* be local min^m & Lagrange multiplier

$$\nabla_x \mathcal{L} = \begin{bmatrix} 1 + 2\lambda x_1 \\ 1 + 2\lambda x_2 \end{bmatrix} = 0$$

$$\Rightarrow x_1 = -\frac{1}{2\lambda} = x_2$$

$$\nabla_\lambda \mathcal{L} = x_1^2 + x_2^2 - 2 = 0$$

$$\Rightarrow \frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = 2$$

$$\Rightarrow 4\lambda^2 = 1$$

$$\Rightarrow \lambda = +1/2 \quad \text{or} \quad -1/2$$

$x_1 = x_2 = -1$ $\qquad$ $x_1 = x_2 = 1$

Candidates for optimality $\to x_1 = 1, x_2 = 1$

$\to x_1 = -1, x_2 = -1$

At $(1,1)$ $\quad x_1 + x_2 = 2$

At $(-1,-1)$ $\quad x_1 + x_2 = -2$

$\therefore x_1 = 1 \; x_2 = 1$ is the unique global max^m

& $x_1 = -1 \; x_2 = -1$ " " " " min^m

They are unique $\because$ only 2 candidate pts.

PS 1 Problem 3

Thursday, April 4, 2019 5:13 PM

$$\min \quad \frac{1}{2}(x_1^2 + x_2^2 + x_3^2)$$

$$\text{st} \quad x_1 + x_2 + x_3 \le -3$$

$$\mathcal{L} = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2) + \mu(x_1 + x_2 + x_3 + 3)$$

By KKT NOC, if x^* is a local minima

$\Rightarrow \nabla_x \mathcal{L} = 0$ & $\exists\ \mu_j^* \geq 0 \quad \forall j \in A(x^*)$

$\mu_j^* = 0 \quad \forall j \notin A(x^*)$

Case 1 $x_1 + x_2 + x_3 + 3 \le 0$ active

$\Rightarrow \nabla_x \mathcal{L} = 0$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0$$

$\Rightarrow x_1 = -\mu = x_2 = x_3$

$x_1 + x_2 + x_3 + 3 = 0$

$\Rightarrow 3 - 3\mu = 0$

$\Rightarrow \mu = 1 \quad \geq 0$ ✓

$$\therefore \left.\begin{array}{l} x_1 = -1 \\ x_2 = -1 \\ x_3 = -1. \end{array}\right\} \text{is a candidate for optimality}$$

Case 2 constraint inactive $\Rightarrow \mu = 0$

$\Rightarrow \nabla_x \mathcal{L} = 0$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\Rightarrow \left.\begin{array}{l} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{array}\right\} \text{also a candidate}$$

$$\boxed{\text{Candidates: } \begin{array}{l} x_1=-1 \\ x_2=-1 \\ x_3=-1 \end{array} \ \& \ \begin{array}{l} x_1=0 \\ x_2=0 \\ x_3=0 \end{array}}$$

We see that $\frac{1}{2}(x_1^2+x_2^2+x_3^2)$ minimized at $x_1=0=x_2=x_3$. So $(0,0,0)$ is local min^m.

PS 1 Problem 4

Thursday, April 4, 2019 5:23 PM

$$\min_{x \in \mathbb{R}^n} \quad f(x) := \frac{1}{2} x^T Q x - b^T x$$

$Q \in \mathbb{R}^{n \times n}$ pos def $\lambda_1, \dots, \lambda_n > 0$ (symm)

$b \in \mathbb{R}^n$

a) candidate x^* for local min^m

$\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ pos def

$\Rightarrow Q x^* - b = 0$ $\Rightarrow Q$ pos def

$\Rightarrow \boxed{x^* = Q^{-1} b}$ ✓ given!

$\therefore$ only candidate is $x^* = Q^{-1} b$

$\because$ Q is pos def everywhere

$\Rightarrow \nabla^2 f(x)$ " " " "

$\Rightarrow f$ is strictly convex

$\therefore$ $\boxed{x^* = Q^{-1} b \text{ is both local \& global min}^m}$

b) After $x^{(0)} \in \mathbb{R}^n$, we pick

$$x^{(1)} = x^{(0)} - \alpha^k \left(\nabla^2 f(x^{(0)})\right)^{-1} \nabla f(x^{(0)})$$

If we pick step size $\alpha = 1 = \eta_0$

$$x^{(1)} = x^{(0)} - 1 \; Q^{-1} (Q x^{(0)} - b)$$

$$\Rightarrow x^{(1)} = x^{(0)} - x^{(0)} + Q^{-1} b$$

$$\Rightarrow \boxed{x^{(1)} = Q^{-1} b = x^*}$$

So, we converge in 1 iteration to x^*

> If n is large and Q has no particular structure, then computing inverse of $Q \in \mathbb{R}^{n \times n}$ is very computationally expensive, making this method intractable.

c) $S \in \mathbb{R}^{n\times n}$ is symm

$\Rightarrow S = U \Sigma U^T$ $\quad U \in \mathbb{R}^{n\times n}$ orthogonal

$\Sigma = diag(\mu_1, \dots \mu_n)$

$x \in \mathbb{R}^n$

$\|Sx\|_2 = \|U \Sigma U^T x\|_2$

Let $\Sigma U^T x = y$ $\quad \Rightarrow y \in \mathbb{R}^{n\times n}\ \mathbb{R}^{n\times n}\ \mathbb{R}^{n\times 1}$

$\Rightarrow y \in \mathbb{R}^n$

$\therefore \|Uy\|_2 = \|U^T y\|_2 = \|y\|_2$

$\because U$ is orthogonal

$$\boxed{\therefore \|Sx\|_2 = \|U \underbrace{\Sigma U^T x}_{y}\|_2 = \|\Sigma U^T x\|_2}$$

For any $z \in \mathbb{R}^n$

$$\Sigma z = \begin{bmatrix} \mu_1 & & 0 \\ & \mu_2 \ddots & \\ 0 & & \mu_n \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} \mu_1 z_1 \\ \mu_2 z_2 \\ \vdots \\ \mu_n z_n \end{bmatrix}$$

Let $|\mu^*| = \max_{i=1:n} |\mu_i|$

$$\Rightarrow \|\Sigma z\| = \left\| \begin{bmatrix} \mu_1 z_1 \\ \vdots \\ \mu_n z_n \end{bmatrix} \right\| = \left(\sum_{i=1:n} \mu_i^2 z_i^2 \right)^{1/2}$$

Each $\mu_i^2 z_i^2 \leq \mu^{*2} z_i^2$ $\quad \because |\mu^*| \geq |\mu_i|$ by defn

$$\Rightarrow \|\Sigma z\|_2 \leq \left(\mu^{*2} \sum_{i=1:n} z_i^2 \right)^{1/2} = \mu^* \|z\|_2$$

$$\boxed{\Rightarrow \|\Sigma z\|_2 \leq \left(\max_{i=1:n} |\mu_i| \right) \|z\|_2}$$

We showed $\|Sx\|_2 = \|\Sigma \underbrace{U^T x}_{z}\|_2$

Let $U^T x = z$

$$\|\Sigma z\|_2 \leq \left(\max_{i=1:n} |\mu_i| \right) \|z\|_2$$

$$\Rightarrow \|Sx\|_2 \leq \left(\max_{i=1:n} |\mu_i| \right) \|U^T x\|_2$$

$\because U$ orthogonal $\|U^T x\|_2 = \|x\|_2$

$$\Rightarrow \boxed{\|Sx\|_2 \le \left(\max_{i=1:n} |\mu_i|\right) \|x\|_2}$$

d) $\eta > 0$ eig of $Q \to \lambda_1 \dots \lambda_n$ $\|Q - \lambda I\| = 0$

$$\Rightarrow Q v_i = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} v_i$$

eig of $I - \eta Q$. Choose v_i = eigvectors of Q.

$$(I - \eta Q) v_i = I v_i - \eta Q v_i$$
$$= v_i - \eta \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} v_i$$
$$= \begin{bmatrix} 1-\eta\lambda_1 & & 0 \\ & \ddots & \\ 0 & & 1-\eta\lambda_n \end{bmatrix} v_i$$

$\therefore$ $\boxed{(1-\eta\lambda_1), (1-\eta\lambda_2) \text{ ----- } (1-\eta\lambda_n) \text{ are eigenvalues of } (I - \eta Q).}$

e) $x^{(k+1)} = x^{(k)} - \eta \nabla f(x^{(k)})$

$\delta_{k+1} := \|x^{(k+1)} - x^*\|_2$ $\delta_k := \|x^{(k)} - x^*\|_2$

$x^* = Q^{-1} b \Rightarrow b = Q x^*$

$\delta_{k+1} = \|x^{(k)} - \eta \nabla f(x^{(k)}) - x^*\|_2$

$\cdot$ $f(x) = \frac{1}{2} x^T Q x - b^T x$

$\nabla f(x^{(k)}) = Q x^{(k)} - b$

$$\delta_{k+1} = \|x^{(k)} - \eta Q x^{(k)} + \eta b - x^*\|_2$$
$$= \|x^{(k)} - x^* - \eta Q x^{(k)} + \eta Q x^*\|_2$$
$$= \|\underbrace{(I - \eta Q)}_{S} \underbrace{(x^{(k)} - x^*)}_{y}\|_2$$

We showed $\|Sy\|_2 \le \left(\max_{i=1:n} |\mu_i|\right) \|y\|_2$

where μ_i are eigenvalues of S

Here eigenvalues of $I - \eta Q$ are $1 - \eta \lambda_i$ where λ_i are eigenvalues of Q

$$\therefore \delta_{k+1} = \|(I - \eta Q)(x^{(k)} - x^*)\|_2$$

$$\leq \left(\max_{i=1:n} |1 - \eta \lambda_i|\right) \|x^{(k)} - x^*\|_2$$

$$\Rightarrow \boxed{\begin{array}{l} \delta_{k+1} \leq \gamma(\eta)\, \delta_k \\ \text{where } \gamma(\eta) = \max_{i=1:n} |1 - \eta \lambda_i| \end{array}}$$

By induction,

$$\delta_1 \leq \gamma(\eta)\, \delta_0$$

$$\delta_k \leq \gamma(\eta)\, \delta_{k-1} \leq \gamma(\eta)^2\, \delta_{k-2} \leq \gamma(\eta)^3 \delta_{k-3} \dots$$

$$\Rightarrow \boxed{\delta_k \leq \gamma(\eta)^k\, \delta_0}$$

We want $\lim_{k \to \infty} x^{(k)} = x^*$

$$\Rightarrow \lim_{k \to \infty} \delta_k = 0$$

$$\Rightarrow \text{req'd } 0 < \gamma(\eta) < 1$$

$$\Rightarrow \max_{i=1:n} |1 - \eta \lambda_i| < 1$$

$$\Rightarrow \forall\, i = 1:n \quad 0 < \eta \lambda_i < 2$$

$$\Rightarrow \forall\, i = 1:n \quad 0 < \eta < 2/\lambda_i$$

f) $d_k = -\nabla f(x^{(k)})$

$$\eta_k = \arg\min_{\eta \geq 0} f(x^{(k)} + \eta d_k)$$

$$f(x^{(k)}) = \frac{1}{2} x^{(k)T} Q x^{(k)} - b^T x^{(k)}$$

$$\nabla f(x^{(k)}) = Q x^{(k)} - b$$

$$\Rightarrow d_k = -Q x^{(k)} + b \qquad \Rightarrow x^{(k)} = Q^{-1}(b - d_k)$$

η_k st $f(x^{(k)} + \eta d_k)$ minimized wrt η

$$x^{(k)} + \eta d_k = Q^{-1} b - Q^{-1} d_k + \eta d_k$$

$$\therefore f(x^{(k)} + \eta d_k) = \frac{1}{2} (Q^{-1} b - Q^{-1} d_k + \eta d_k)^T Q (Q^{-1} b - Q^{-1} d_k + \eta d_k) - b^T (Q^{-1} b - Q^{-1} d_k + \eta d_k)$$

$$\therefore f(x^k + \eta d_k) = \frac{1}{2}(Q^{-1}b - Q^{-1}d_k + \eta d_k)^T Q (Q^{-1}b - Q^{-1}d_k + \eta d_k) \ldots$$

$$\frac{\partial f(x^k + \eta d_k)}{\partial \eta} = 0$$

$$\Rightarrow d_k^T Q (Q^{-1} b - Q^{-1} d_k + \eta d_k) \quad - d_k^T b = 0$$

$$\Rightarrow \cancel{d_k^T b} - d_k^T d_k + \eta d_k^T Q \cdot d_k - \cancel{d_k^T b} = 0$$

$$\Rightarrow \eta \, d_k^T Q d_k = \|d_k\|_2^2$$

$$\Rightarrow \boxed{\eta_k = \frac{\|d_k\|_2^2}{d_k^T Q d_k}}$$

g) $x^* = Q^{-1} b \qquad n = 2 \quad \gamma = 10$

$f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2) \qquad Q : \mathbb{R}^{2\times 2}$

$$= \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \delta & \varepsilon \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} \alpha x_1 + \delta x_2 & \beta x_1 + \varepsilon x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - (b_1 x_1 + b_2 x_2)$$

$$= \frac{1}{2}\left(\alpha x_1^2 + (\delta + \beta) x_1 x_2 + \varepsilon x_2^2\right) - (b_1 x_1 + b_2 x_2)$$

$\Rightarrow \alpha = 1 \qquad \varepsilon = \gamma \qquad \beta = \delta = 0$

$b_1 = b_2 = 0$

$$\Rightarrow x^* = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla f(x^{(k)}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} x_1 \\ \gamma x_2 \end{bmatrix}$$

$\lambda_i = \text{eig}(Q) = 1, \ \gamma$

reqd for const step size $0 < \eta < 2/\lambda_i \ \forall i$

$\Rightarrow 0 < \eta < 2$ and

$0 < \eta < 2/\gamma$

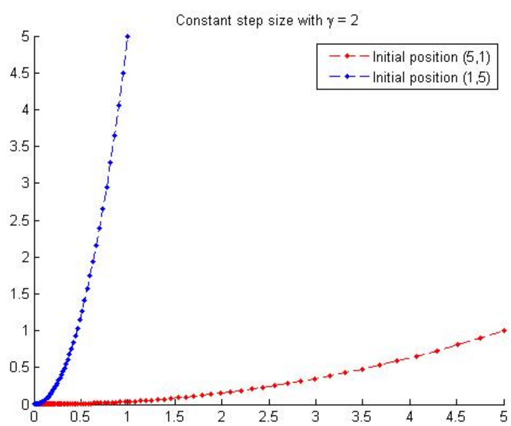
So we pick $\eta < 2/10$

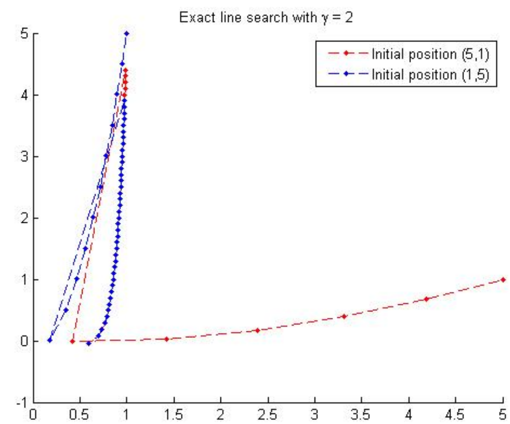
$\eta = 0.05$ for example

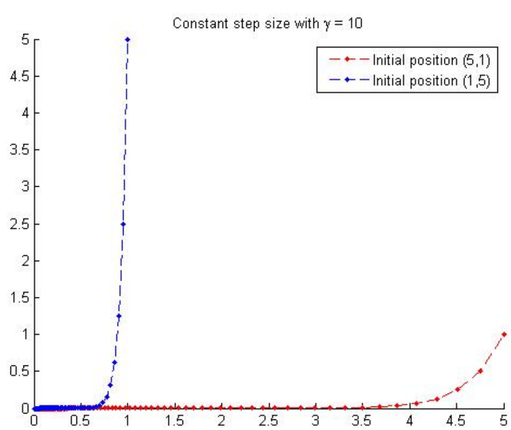
Optimal soln: $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

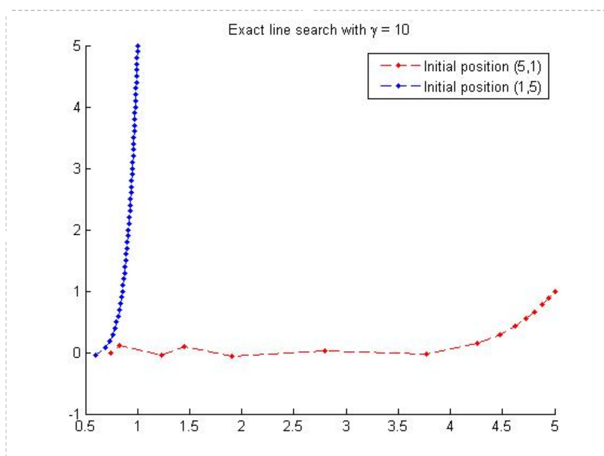
Exact line search:
- finds soln faster
- for lower $\gamma = 2$, it overshoots the solution
- we see a zig zag close to the optimal solution where slight (+)ve or (−)ve gradients cause zig zagging due to sharp change in η

Constant step size:
- takes many iterations to find solution
- takes a smooth path to origin for both $\gamma = 10$ and $\gamma = 2$









PS 1 Problem 5

Tuesday, April 9, 2019 4:42 PM

$x_{t+1} = Ax_t + Bu_t$ LTI

$$\min_{u \in \mathbb{R}^{mT}} J(u) := x_T^T Q_T x_T + \sum_{t=0}^{T-1} x_t^T Q x_t + u_t^T R u_t$$

equivalent to

$$\min_{u \in \mathbb{R}^{mT}} \frac{1}{2} u^T \tilde{Q} u - \tilde{b}^T u$$

$x_1 = Ax_0 + Bu_0$

$x_2 = A(Ax_0 + Bu_0) + Bu_1$

$= A^2 x_0 + ABu_0 + Bu_1$

$x_3 = A^3 x_0 + A^2 B u_0 + ABu_1 + Bu_2$

$$\Rightarrow x_T = A^T x_0 + \sum_{i=0}^{T-1} A^{T-1-i} B u_i$$

$$x_t = A^t x_0 + \sum_{i=0}^{t-1} A^{t-1-i} B u_i$$

$$x_t^T Q x_t + u_t^T R u_t = \left(A^t x_0 + \sum_{i=0}^{t-1} A^{t-1-i} B u_i\right)^T Q \left(A^t x_0 + \sum_{i=0}^{t-1} A^{t-1-i} B u_i\right) + u_t^T R u_t$$

$$\because \left(\sum_{i=0}^{t-1} A^{t-1-i} B u_i\right)' Q A^t x_0 \text{ is a scalar}$$

$$= (A^t x_0)' Q \left(\sum_{i=0}^{t-1} A^{t-1-i} B u_i\right)$$

$$\therefore x_t' Q x_t + u_t' R u_t = x_0' A^{t'} Q A^t x_0 + \left(\sum_{i=0}^{t-1} A^{t-1-i} B u_i\right)' Q \left(\sum_{i=0}^{t-1} A^{t-1-i} B u_i\right) + 2 x_0' A^{t'} Q \left(\sum_{i=0}^{t-1} A^{t-1-i} B u_i\right) + u_t' R u_t$$

We can write

$$\left(\sum_{i=0}^{t-1} A^{t-1-i} B u_i\right) = \begin{bmatrix} A^{t-1}B & & & 0 \\ & A^{t-2}B & & \\ & & \ddots & AB \\ 0 & & & B \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_{t-1} \end{bmatrix}$$

$$= C(t)\, u_{0:t-1}$$

$C(t) \in \mathbb{R}^{mt \times mt}$ $\quad u_{0:t-1} \in \mathbb{R}^{mt}$

$$\therefore x_t' Q x_t + u_t' R u_t = x_0' A^{t'} Q A^t x_0 + u_{0:t-1}' C(t)' Q C(t) u_{0:t-1} + 2 x_0' A^{t'} Q C(t) u_{0:t-1} + u_t' R u_t$$

$$\sum_{t=0}^{T-1} x_t' Q x_t + u_t' R u_t = \underbrace{\sum_{t=0}^{T-1} x_0' A^{t'} Q A^t x_0}_{1} + \underbrace{\sum_{t=0}^{T-1} u_{0:t-1}' C(t)' Q C(t) u_{0:t-1}}_{2} + \underbrace{\sum_{t=0}^{T-1} 2 x_0' A^{t'} Q C(t) u_{0:t-1}}_{3} + \underbrace{\sum_{t=0}^{T-1} u_t' R u_t}_{4}$$

$$\therefore x_T = \overset{\text{power } T}{A^T} x_0 + \sum_{i=0}^{T-1} A^{T-1-i} B u_i = \overset{\text{power } T}{A^T} x_0 + C(T)\, u_{0:T-1}$$

$$J(u) = x_T' Q_T x_T + ① + ② + ③ + ④$$

$$= \underbrace{x_0'(A^T)' Q_T A^T x_0}_{⑤} + \underbrace{u_{0:T-1}' C(T)' Q_T C(T) u_{0:T-1}}_{⑥}$$

$$+ \underbrace{2 x_0'(A^T)' Q_T C(T) u_{0:T-1}}_{⑦} + ① + ② + ③ + ④$$

① and ⑤ are constant wrt u and are only functions of x_0 so we can leave them out of $\min_{u \in \mathbb{R}^{mT}} J(u)$

③ and ⑦ $\quad \sum_{t=0}^{T-1} 2x_0' A^{t'} Q C(t) u_{0:t-1} + 2 x_0'(A^T)' Q_T C(T) u_{0:T-1}$

$$= 2x_0' \Big[\left(A^{0'} Q A^0 B + A^{1'} Q A^1 B + \dots + A^{T-1'} Q A^{T-1} B\right) u_0$$
$$+ \left(A^{1'} Q A^0 B + A^{2'} Q A^1 B + \dots + A^{T-1'} Q A^{T-2} B\right) u_1$$
$$+ \left(A^{2'} Q A^0 B + A^{3'} Q A^1 B + \dots + A^{T-1'} Q A^{T-3} B\right) u_2$$
$$+ \dots$$
$$+ \left(A^{T-1'} Q A^0 B\right) u_{T-1} \Big]$$

$$+ 2x_0'(A^T)' Q_T \left[A^{T-1} B u_0 + A^{T-2} B u_1 + \dots A^0 B u_{T-1} \right]$$

$$= 2x_0' \Big[\sum_{i=0}^{T-1} (A^i)' Q A^i B + (A^T)' Q_T A^{T-1} B, \; \sum_{i=1}^{T-1} (A^i)' Q A^{i-1} B + (A^T)' Q_T A^{T-2} B, \; \dots \; \sum_{i=T-1}^{T-1} (A^i)' Q A^{i-(T-1)} B + (A^T)' Q_T \dots$$

$$= -2 \tilde{b}^T u$$

$$\boxed{\text{where } \tilde{b} = \begin{bmatrix} \tilde{b}_0 \\ \tilde{b}_1 \\ \vdots \\ \tilde{b}_{T-1} \end{bmatrix} \quad \tilde{b}_j = -x_0' \left(\sum_{i=j}^{T-1} (A^i)' Q A^i B + (A^T)' Q_T A^{T-1-j} B \right)}$$

Check $\tilde{b}_j$: $1\times n \quad n\times n \quad n\times n \quad n\times n \; n\times m \quad = 1\times m$

$$\Rightarrow \tilde{b} : 1\times mT \quad \Rightarrow \tilde{b}^T : \mathbb{R}^{mT}$$

② ④ ⑥

$$= \sum_{t=0}^{T-1} u_{0:t-1}' C(t)' Q\, C(t)\, u_{0:t-1} + \sum_{t=0}^{T-1} u_t' R u_t + u_{0:T-1}' C(T)' Q_T\, C(T)\, u_{0:T-1}$$

$$= u' \tilde{Q} u$$

$$C(t)\, u_{0:t-1} = \sum_{i=0}^{t-1} A^{t-1-i} B u_i$$

$C(t)$ is a diagonal matrix.

$$\boxed{\begin{gathered} \tilde{Q} = \begin{bmatrix} \tilde{Q}_{00} & & & 0 \\ & \tilde{Q}_{11} & & \\ & & \ddots & \\ 0 & & & \tilde{Q}_{(T-1)(T-1)} \end{bmatrix} \\ \tilde{Q}_{jj} = \sum_{i=j}^{T-1} B' A^{i\,\prime} Q A^{i} B + R + B' A^{T-1-j\,\prime} Q_T A^{T-1-j} B \end{gathered}}$$

Check $\tilde{Q}_{jj}$: $m\times n \quad n\times n \quad n\times n \quad n\times n \quad n\times m$

$= m\times m$

$\Rightarrow \tilde{Q} : mT \times mT$

b) $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad Q^T = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \quad R = 1$

$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad x_0 = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad T = 20$

$\Rightarrow n = 2 \quad m = 1$

$\tilde{b} \in \mathbb{R}^{mT} = \mathbb{R}^{20}$

$$\tilde{b}_0 = -x_0' \left(\sum_{i=0}^{19} (A^i)' Q A^i B + (A^{20})' Q_T A^{19} B \right)$$

$\tilde{Q} : mT \times mT = 20 \times 20$

$$\tilde{Q}_{00} = \sum_{i=0}^{19} B' A^{19-i\,\prime} Q A^{19-i} B + R + B' A^{19\,\prime} Q_T A^{19} B$$

$$A^2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$$

Solving $u^* = \tilde{Q}^{-1} \tilde{b}$

```
uStar =

   -0.0622
   -0.0645
   -0.0666
   -0.0687
   -0.0706
   -0.0723
   -0.0738
   -0.0752
   -0.0763
   -0.0773
   -0.0779
   -0.0783
   -0.0784
   -0.0783
   -0.0777
   -0.0766
   -0.0749
   -0.0720
   -0.0664
   -0.0509
```

cost

```
J =

   2.0933e+03
```

AA 203 HW 1 Question 5

Somrita Banerjee

```
clc
clear all
close all
Q = eye(2);
QT = 10 * eye(2);
R = eye(1);
A=[1 1; 0 1];
B=[0;1];
x0=[1;0];
T=20;
btilde = zeros(T,1);
Qtilde = zeros(T,T);
for j = 0: T-1
    sumb = 0;
    sumQ = 0;
    for i = j: T-1
        sumb = sumb + (-x0' * (A^i)' * Q * (A^i) * B);
        sumQ = sumQ + B' * (A^i)' * Q * (A^i) *B;
    end
    btilde(j+1,1) =  sumb - x0'*((A^T)'*QT*(A^(T-1-j))*B);
    Qtilde(j+1,j+1) = sumQ + R + B' * (A^(T-1-j))' * QT * (A^(T-1-j)) * B;
end
uStar = inv(Qtilde)*btilde;
x = zeros(2, 21);
x(:,1) = [1; 0];
i = 1;
u = uStar;

sumJ = 0;
for t = 0:T-1
    x(:,t+2) = A*x(:,t+1) + B*u(t+1);
    sumJ = sumJ + x(:,t+1)'*Q*x(:,t+1) + u(t+1)'*R*u(t+1);
end
J = x(:,T+1)'*QT*x(:,T+1) + sumJ

QuadCost = 0.5*u'*Qtilde*u - btilde'*u
```

```
J =

   2.0933e+03


QuadCost =

 -164.2338
```

Published with MATLAB® R2014a