# 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<br>($\nabla f_1 = 0$ $\nabla^2 f_1$ pos semi-def) for local min<br>$(0,0)$ also satisfies SOC<br>($\nabla f_1 = 0$ $\nabla^2 f_1$ 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

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

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 $x^*, \lambda^*$ be local min$^m$ & Lagrange multiplier

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

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

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

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

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

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

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

Candidates for optimality $\longrightarrow$ $x_1 = 1,\ x_2 = 1$
$\longrightarrow$ $x_1 = -1,\ x_2 = -1$

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

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

$$\boxed{\begin{array}{l} \therefore\ x_1 = 1 \quad x_2 = 1 \quad \text{is the unique global max}^m \\ \&\ x_1 = -1 \quad x_2 = -1 \quad \text{" " " " min}^m \end{array}}$$

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

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

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

<u>Case 1</u> $x_1 + x_2 + x_3 + 3 \le 0$ active

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

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

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

$x_1 + x_2 + x_3 + 3 = 0$

$\Rightarrow 3 - 3\mu = 0$

$\Rightarrow \mu = 1 \quad \ge 0$ ✓

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

<u>Case 2</u> constraint inactive $\Rightarrow \mu = 0$

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

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

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

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

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 $\lambda_1, \ldots \lambda_n > 0$ (symm)

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

a) candidate $x^*$ for local min$^m$

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

$\Rightarrow Qx^* - b = 0$ $\Rightarrow Q$ 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

$\Sigma = \text{diag}(\mu_1, \dots \mu_n)$

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

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

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

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

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

$\because U$ is orthogonal

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

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

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

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

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

Let $U^T x = z$

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

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

$\because U$ 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 $Q \rightarrow \lambda_1 \dots \lambda_n$ $\qquad \|Q - \lambda I\| = 0$

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

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

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

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

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

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

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

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

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

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

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

where $\mu_i$ are eigenvalues of $S$

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

$$\therefore\ \delta_{k+1} = \left\|(I-\eta Q)(x^{(k)}-x^*)\right\|_2$$
$$\le \left(\max_{i=1:n} |1-\eta\lambda_i|\right) \left\|x^{(k)}-x^*\right\|_2$$

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

By induction,

$$\delta_1 \le \gamma(\eta)\,\delta_0$$
$$\delta_k \le \gamma(\eta)\,\delta_{k-1} \le \gamma(\eta)^2\,\delta_{k-2} \le \gamma(\eta)^3\,\delta_{k-3} \cdots$$
$$\Rightarrow \boxed{\delta_k \le \gamma(\eta)^k\,\delta_0}$$

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

$$\Rightarrow \lim_{k\to\infty} \delta_k = 0$$
$$\Rightarrow \text{reqd } 0 < \gamma(\eta) < 1$$
$$\Rightarrow \max_{i=1:n} |1-\eta\lambda_i| < 1$$
$$\Rightarrow \forall\, i=1:n \quad 0 < \eta\lambda_i < 2$$
$$\Rightarrow \forall\, i=1:n \quad 0 < \eta < 2/\lambda_i$$

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

$$\eta_k = \operatorname*{argmin}_{\eta\ge 0} f(x^{(k)}+\eta d_k)$$
$$f(x^{(k)}) = \tfrac{1}{2} x^{(k)T} Q\, x^{(k)} - b^T x^{(k)}$$
$$\nabla f(x^{(k)}) = Q x^{(k)} - b$$
$$\Rightarrow d_k = -Q x^{(k)} + b \qquad \Rightarrow x^{(k)} = Q^{-1}(b-d_k)$$

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

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

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

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

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

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

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

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

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

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

$0 < \eta < 2/\gamma$

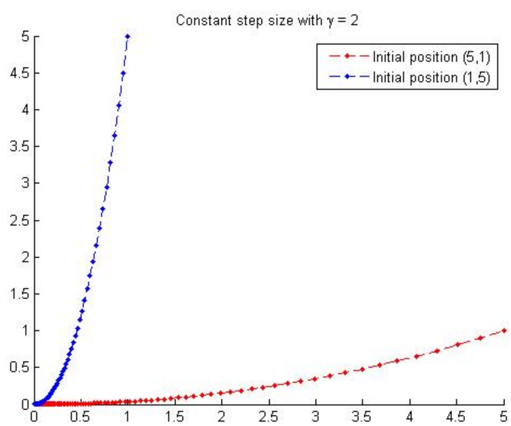
So we pick $\eta < 2/10$

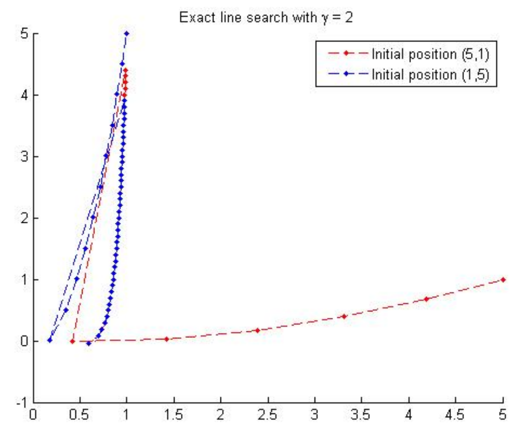
$\eta = 0.05$ for example

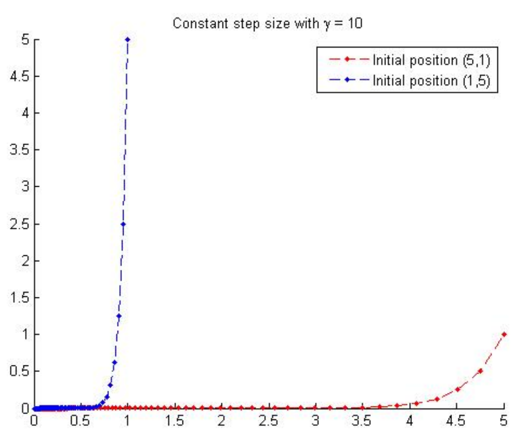
Optimal soln : $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

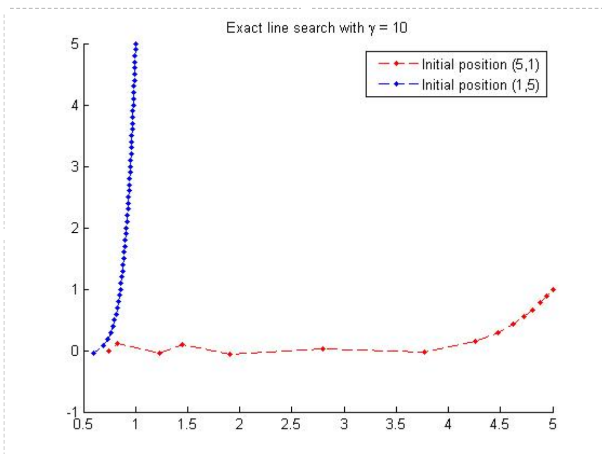
Exact line search:
- finds soln faster
- for lower $\gamma = 2$, it overshoots the solution
- we see a zig zag close to the optimal solution where slight (+)ve or (−)ve gradients cause zig zagging due to sharp change in $\eta$

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

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

$x_1 = Ax_0 + Bu_0$

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

$= A^2 x_0 + ABu_0 + Bu_1$

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

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

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

$A \in \mathbb{R}^{n \times n}$ $B \in \mathbb{R}^{n \times m}$

$$x = \begin{bmatrix} I_{n\times n} \\ A \\ \vdots \\ A^T \end{bmatrix} x_0 + \begin{bmatrix} 0_{n\times m} & 0 & \cdots & 0 \\ B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & 0 & & \\ A^{T-1}B & A^{T-2}B & \cdots & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{T-1} \end{bmatrix}$$

$= \tilde{A} x_0 + \tilde{B} u$

$x_0 \in \mathbb{R}^n$ $\tilde{A} \in \mathbb{R}^{(T+1)n \times n}$

$u \in \mathbb{R}^{Tm}$ $\tilde{B} \in \mathbb{R}^{(T+1)n \times mT}$

$x \in \mathbb{R}^{(T+1)n}$

$$J = x_T' Q_T x_T + \sum_{t=0}^{T-1} x_t' Q x_t + u_t' R u_t$$

$$= x_0' Q x_0 + x_1' Q x_1 + \cdots x_{T-1}' Q x_{T-1} + x_T' Q_T x_T$$
$$+ u_0' R u_0 + u_1' R u_1 + \cdots u_{T-1}' R u_{T-1}$$

$$= x' \underbrace{\begin{bmatrix} Q & & & \\ & Q & & \\ & & \ddots & \\ & & Q & \\ & & & Q_T \end{bmatrix}}_{\hat{Q} \in \mathbb{R}^{(T+1)n \times (T+1)n}} x + u' \underbrace{\begin{bmatrix} R & & \\ & R & \\ & & \ddots \\ & & & R \end{bmatrix}}_{\tilde{R} \in \mathbb{R}^{Tm \times Tm}} u$$

$$= (\tilde{A} x_0 + \tilde{B} u)' \hat{Q} (\tilde{A} x_0 + \tilde{B} u) + u' \tilde{R} u$$

$$= \underbrace{x_0' \tilde{A}' \hat{Q} \tilde{A} x_0}_{\text{const wrt } u} + u' \tilde{B}' \hat{Q} \tilde{B} u + 2 x_0' \tilde{A}' \hat{Q} \tilde{B} u$$
$$+ u' \tilde{R} u$$

$$\Rightarrow \min J = \min \frac{1}{2} u' \underbrace{\left( \tilde{B}' \hat{Q} \tilde{B} + \tilde{R} \right)}_{\tilde{Q}} u + \underbrace{x_0' \tilde{A}' \hat{Q} \tilde{B}}_{-\tilde{b}'} u$$

$$\tilde{Q} = \tilde{B}' \hat{Q} \tilde{B} + \tilde{R}$$

$$\hat{Q} \in \mathbb{R}^{(T+1)n \times (T+1)n} = \begin{bmatrix} Q & & & & \\ & Q & & & \\ & & \ddots & & \\ & & & Q & \\ & & & & Q_T \end{bmatrix}$$

$$\tilde{B} \in \mathbb{R}^{(T+1)n \times mT} = \begin{bmatrix} 0 & & & \\ B & & & \\ AB & B & & \\ \vdots & & & \\ A^{T-1}B & A^{T-2}B & \cdots & B \end{bmatrix}$$

$$\tilde{R} \in \mathbb{R}^{mT \times mT} = \begin{bmatrix} R & & \\ & \ddots & \\ & & R \end{bmatrix}$$

$$\Rightarrow \tilde{Q} \in \mathbb{R}^{mT \times mT}$$

$$\tilde{b} = -\left( x_0' \tilde{A}' \hat{Q} \tilde{B} \right)'$$

$x_0 \in \mathbb{R}^n$ $\quad \tilde{A} \in \mathbb{R}^{(T+1)n \times n}$ $\quad \tilde{A} = \begin{bmatrix} I_{n \times n} \\ A \\ \vdots \\ A^T \end{bmatrix}$

$\hat{Q} \in \mathbb{R}^{(T+1)n \times (T+1)n}$

$\tilde{B} \in \mathbb{R}^{(T+1)n \times mT}$

$\tilde{b} \in \left( \mathbb{R}^{1 \times mT} \right)' \in \mathbb{R}^{mT}$

$$u^* = \tilde{Q}^{-1} \tilde{b} \qquad J(u^*) = 2.9471$$

## AA 203 HW 1 Question 5 again

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);
Qhat = blkdiag(kron(eye(20),Q),QT);
Atilde=eye(2);
for i = 1:T
    Atilde=[Atilde;A^i];
end
```

```
end
Btilde = zeros((T+1)*2,T);
for i=1:T
    for j=1:i
        Btilde(2*i+1: 2*i+2,j)=(A^(i-j)) *B;
    end
end
Rtilde = kron(eye(20),R);
Qtilde = Btilde'*Qhat*Btilde + Rtilde;
btilde = -(x0'*Atilde'*Qhat*Btilde)';
uStar = Qtilde\btilde

u= uStar;
x = zeros(2, T+1);
x(:,1) = x0;
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
```

```
uStar =

   -0.4221
    0.1030
    0.1530
    0.0974
    0.0464
    0.0177
    0.0051
    0.0007
   -0.0004
   -0.0004
   -0.0002
   -0.0001
   -0.0000
   -0.0000
   -0.0000
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000


J =

    2.9471
```