HW3 ECE 719 Optimal systems

Spring 2016
Electrical engineering department
University of Wisconsin, Madison

Instructor: Professor B Ross Barmish

By

Nasser M. Abbasi

December 30, 2019

Contents

List of Figures

List of Tables

0.1 Problem 1

PROBLEM DESCRIPTION

Barmish

ECE 719 – Homework Hyperplane

Given a continuously differentiable convex function J and any pair of points u^1, u^2 in $\mathbf{R}^n$, prove that the inequality

$$J(u^2) \geq J(u^1) + [\nabla J(u^1)]^T (u^2 - u^1)$$

must hold.

SOLUTION

Since $J(\boldsymbol{u})$ is a convex function $J : \Re^n \to \Re$, then by definition of convex functions we write

$$J\left((1-\lambda)\boldsymbol{u}^1 + \lambda \boldsymbol{u}^2\right) \leq (1-\lambda) J\left(\boldsymbol{u}^1\right) + \lambda J\left(\boldsymbol{u}^2\right)$$

Where $\lambda \in (0,1)$. Rewriting the above as follows

$$J\left(\boldsymbol{u}^1 - \lambda \boldsymbol{u}^1 + \lambda \boldsymbol{u}^2\right) \leq J\left(\boldsymbol{u}^1\right) - \lambda J\left(\boldsymbol{u}^1\right) + \lambda J\left(\boldsymbol{u}^2\right)$$

$$J\left(\boldsymbol{u}^1 + \lambda\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)\right) - J\left(\boldsymbol{u}^1\right) \leq \lambda\left(J\left(\boldsymbol{u}^2\right) - J\left(\boldsymbol{u}^1\right)\right)$$

Dividing both sides by $\lambda \neq 0$ gives

$$\frac{J\left(\boldsymbol{u}^1 + \lambda\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)\right) - J\left(\boldsymbol{u}^1\right)}{\lambda} \leq J\left(\boldsymbol{u}^2\right) - J\left(\boldsymbol{u}^1\right)$$

Taking the limit $\lambda \to 0$ results in

$$\lim_{\lambda\to 0} \frac{J\left(\boldsymbol{u}^1 + \lambda\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)\right) - J\left(\boldsymbol{u}^1\right)}{\lambda} \leq \lim_{\lambda\to 0} J\left(\boldsymbol{u}^2\right) - J\left(\boldsymbol{u}^1\right)$$

But $\lim_{\lambda\to 0} \frac{J\left(\boldsymbol{u}^1+\lambda\left(\boldsymbol{u}^2-\boldsymbol{u}^1\right)\right)-J\left(\boldsymbol{u}^1\right)}{\lambda} = \left.\frac{\partial J(\boldsymbol{u})}{\partial\left(\boldsymbol{u}^2-\boldsymbol{u}^1\right)}\right|_{\boldsymbol{u}^1} = \left[\nabla J\left(\boldsymbol{u}^1\right)\right]^T\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)$ (appendix below shows how this came about). Therefore the above becomes

$$\left[\nabla J\left(\boldsymbol{u}^1\right)\right]^T\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right) \leq J\left(\boldsymbol{u}^2\right) - J\left(\boldsymbol{u}^1\right)$$

$$J\left(\boldsymbol{u}^2\right) \geq J\left(\boldsymbol{u}^1\right) + \left[\nabla J\left(\boldsymbol{u}^1\right)\right]^T\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)$$

QED.

0.1.1 Appendix

More details are given here on why

$$\lim_{\lambda\to 0} \frac{J\left(\boldsymbol{u}^1 + \lambda\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)\right) - J\left(\boldsymbol{u}^1\right)}{\lambda} = \left[\nabla J\left(\boldsymbol{u}^1\right)\right]^T\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)$$

Let $\boldsymbol{u}^2 - \boldsymbol{u}^1 = \boldsymbol{d}$. This is a directional vector, its tail starts at $\boldsymbol{u}^1$ going to tip of $\boldsymbol{u}^2$ point. Evaluating $\lim_{\lambda\to 0} \frac{J\left(\boldsymbol{u}^1+\lambda \boldsymbol{d}\right)-J\left(\boldsymbol{u}^1\right)}{\lambda}$ is the same as saying

$$\begin{aligned}
\left.\frac{\partial J(\boldsymbol{u})}{\partial \boldsymbol{d}}\right|_{\boldsymbol{u}^1} &= \lim_{\lambda\to 0} \frac{J\left(\boldsymbol{u}^1 + \lambda \boldsymbol{d}\right) - J\left(\boldsymbol{u}^1\right)}{\lambda} \\
&= \left.\frac{d}{d\lambda} J\left(\boldsymbol{u}^1 + \lambda \boldsymbol{d}\right)\right|_{\lambda=0}
\end{aligned}$$

Using the chain rule gives

$$\begin{aligned}
\left.\frac{d}{d\lambda} J\left(\boldsymbol{u}^1 + \lambda \boldsymbol{d}\right)\right|_{\lambda=0} &= \left.\left[\nabla J\left(\boldsymbol{u}^1 + \lambda \boldsymbol{d}\right)\right]^T \frac{d}{d\lambda}\left(\boldsymbol{u}^1 + \lambda \boldsymbol{d}\right)\right|_{\lambda=0} \\
&= \left.\left[\nabla J\left(\boldsymbol{u}^1 + \lambda \boldsymbol{d}\right)\right]^T \boldsymbol{d}\right|_{\lambda=0} \\
&= \left[\nabla J\left(\boldsymbol{u}^1\right)\right]^T \boldsymbol{d}
\end{aligned}$$

Replacing $\boldsymbol{u}^2 - \boldsymbol{u}^1 = \boldsymbol{d}$, the above becomes

$$
\begin{aligned}
\lim_{\lambda \to 0} \frac{J\left(\boldsymbol{u}^1 + \lambda\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)\right) - J\left(\boldsymbol{u}^1\right)}{\lambda} &= \left.\frac{\partial J(\boldsymbol{u})}{\partial\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)}\right|_{\boldsymbol{u}^1} \\
&= \left[\nabla J\left(\boldsymbol{u}^1\right)\right]^T\left(\boldsymbol{u}^2 - \boldsymbol{u}^1\right)
\end{aligned}
$$

Where $\nabla J\left(\boldsymbol{u}^1\right)$ is the gradient vector of $J(\boldsymbol{u})$ evaluated at $\boldsymbol{u} = \boldsymbol{u}^1$.

0.2 Problem 2

PROBLEM DESCRIPTION

Barmish

ECE 717 – Homework Eigenvalue

Let $M(q)$ be an $n \times n$ symmetric matrix with entries $M_{i,j}(q)$ which depend convexly on a vector $q \in \mathbf{R}^n$. Show that the largest eigenvalue of $M(q)$, call it $\lambda_{max}(q)$, also depends convexly on q.

SOLUTION

Since each $m_{ij}\left(q\right)$ is convex function in q, then

$$
m_{ij}\left((1-\alpha)q^1 + \alpha q^2\right) \leq (1-\alpha)\, m_{ij}\left(q^1\right) + \alpha m_{ij}\left(q^2\right) \tag{1}
$$

For $\alpha \in [0,1]$. We also know by Rayleigh quotient theorem which applies for symmetric matrices that largest eigenvalue of a symmetric matrix is given by

$$
\lambda_{\max} = \max_{x \in \Re^n, \|x\|=1} x^T M x
$$

Therefore, evaluated at point $q^\alpha = (1-\alpha)\, q^1 + \alpha q^2$, the above become

$$
\lambda_{\max}\left((1-\alpha)q^1 + \alpha q^2\right) = \max_{\|x\|=1} \sum_{i,j}^{n} m_{ij}\left((1-\alpha)q^1 + \alpha q^2\right) x_i x_j \tag{2}
$$

Applying (1) in RHS (2) changes $=$ to $\leq$ giving

$$
\begin{aligned}
\lambda_{\max}\left((1-\alpha)q^1 + \alpha q^2\right) &\leq \max_{\|x\|=1} \sum_{i,j}^{n} \left((1-\alpha)\, m_{ij}\left(q^1\right) + \alpha m_{ij}\left(q^2\right)\right) x_i x_j \\
&= \max_{\|x\|=1} \left(\sum_{i,j}^{n} (1-\alpha)\, m_{ij}\left(q^1\right) x_i x_j + \sum_{i,j}^{n} \alpha m_{ij}\left(q^2\right) x_i x_j \right) \\
&= (1-\alpha)\left(\max_{\|x\|=1} \sum_{i,j}^{n} m_{ij}\left(q^1\right) x_i x_j \right) + \alpha\left(\max_{\|x\|=1} \sum_{i,j}^{n} m_{ij}\left(q^2\right) x_i x_j \right)
\end{aligned} \tag{3}
$$

Since

$$
\max_{\|x\|=1} \sum_{i,j}^{n} m_{ij}\left(q^1\right) x_i x_j = \lambda_{\max}\left(q^1\right)
$$

And

$$
\max_{\|x\|=1} \sum_{i,j}^{n} m_{ij}\left(q^2\right) x_i x_j = \lambda_{\max}\left(q^2\right)
$$

Then (3) becomes

$$
\lambda_{\max}\left((1-\alpha)q^1 + \alpha q^2\right) \leq (1-\alpha)\, \lambda_{\max}\left(q^1\right) + \alpha \lambda_{\max}\left(q^2\right)
$$

This is the definition of convex function, therefore $\lambda_{\max}$ is a convex function in q.

Note: I tried also to reduce this to a problem where I could argue that the pointwise maximum of convex functions is also a convex function to solve it. I could not get a clear way to do this, so I solved it as above. I hope I did not violate the cardinal rule by using $\lambda_{\max} = \max_{x \in \Re^n, \|x\|=1} x^T M x$.

0.3 Problem 3

PROBLEM DESCRIPTION

Barmish

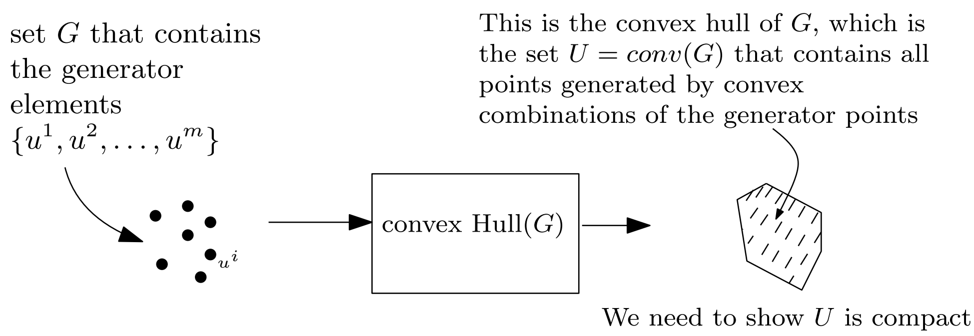
ECE 717 – Homework Polytope

Let U be a polytope in $\mathbf{R}^n$ with generators $u^1, u^2, ..., u^m$. We often describe U by writing

$$U = conv\{u^1, u^2, ..., u^m\}$$

and say the U is the *convex hull* of the u^i. Show that U is compact.

SOLUTION



To show U is bounded, a proof by induction is used. From the definition of constructing U

$$U = \left\{ x \in \Re^n : x = \sum_{i=1}^{m} \lambda_i u^i \right\}$$

Where $\sum_{i=1}^{m} \lambda_i = 1$ and $\lambda_i \geq 0$.

For $m = 1$, $x = \lambda u^1$. So U contains just one element u^1. Since $\lambda = 1$ and u^1 is given and bounded, then this is closed and bounded set with one element. Hence compact. Now we assume U is compact for $m = k - 1$ and we need to show it is compact for $m = k$. In other words, we assume that each $x^* \in U$ generated using

$$x^* = \sum_{i=1}^{k-1} \lambda_i u^i$$

Is such that $\|x^*\| < \infty$ and $x^* \in U$. Now we need to show that U is bounded when generator contains k elements. Now

$$\begin{aligned} x &= \sum_{i=1}^{k} \lambda_i u^i \\ &= \lambda_1 u^1 + \lambda_2 u^2 + \cdots + \lambda_{k-1} u^{k-1} + \lambda_k u^k \end{aligned}$$

Multiply and divide by $(1 - \lambda_k)$

$$\begin{aligned} x &= (1 - \lambda_k) \left(\frac{\lambda_1 u^1}{(1 - \lambda_k)} + \frac{\lambda_2}{(1 - \lambda_k)} u^2 + \cdots + \frac{\lambda_{k-1} u^{k-1}}{(1 - \lambda_k)} + \frac{\lambda_k}{(1 - \lambda_k)} u^k \right) \\ &= (1 - \lambda_k) \left(\sum_{i=1}^{k-1} \frac{\lambda_i}{(1 - \lambda_k)} u^i + \frac{\lambda_k}{(1 - \lambda_k)} u^k \right) \\ &= (1 - \lambda_k) \left(\sum_{i=1}^{k-1} \frac{\lambda_i}{(1 - \lambda_k)} u^i \right) + \lambda_k u^k \end{aligned}$$

But $\sum_{i=1}^{k-1} \frac{\lambda_i}{(1-\lambda_k)} u^i = x^*$ which we assumed in U. Hence the above becomes

$$x = (1 - \lambda_k)\, x^* + \lambda_k u^k$$

Since u^k is element in the generator set G and it is in U by definition, then the above is convex combination of two elements in U. Hence x in also in U (it is on a line between x^* and u^k, both in U). Therefore U is closed and bounded for any m in the generator set. Hence U is compact.

0.4 Problem 4

PROBLEM DESCRIPTION

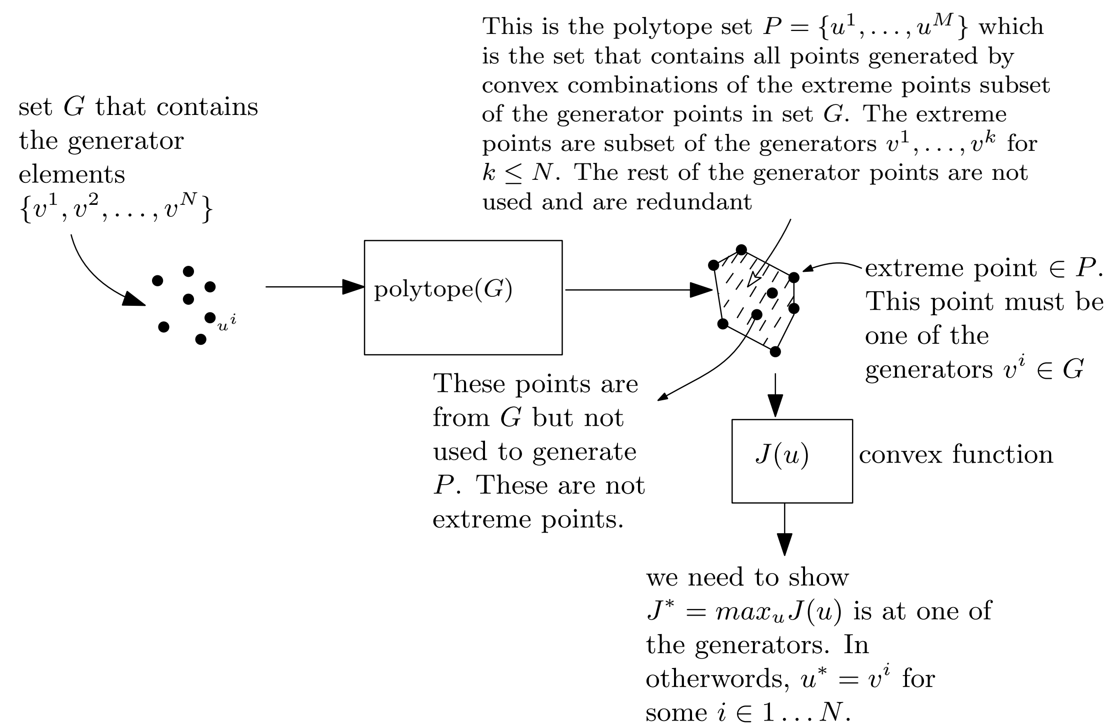
Barmish

ECE 717 – Homework Maximum

Let P be a polytope in $\mathbf{R}^n$ with generators $v^1, v^2, ..., v^N$ and assume $J(u)$ is convex. Prove that the maximum of J subject to $u \in P$ is attained at one of the generators.

Note: this type of result does not hold for the minimum as evidenced by the simple example $J(u) = u^2$ on $[-1, 1]$.

SOLUTION



The extreme points of P are subset of G. They are the points used to generate P. The set P is compact (by problem 3) and convex set (by construction, since it is convex combinations of its extreme points). If we can show that J^* is at an extreme point of P, then we are done, since an extreme point of P is in G.

Let $u^* \in P$ be the point where $J(u)$ is maximum. u^* is a convex combinations of all extreme points of P, (these are also subset from G but they can be the whole set G also if there were no redundant generators), Therefore

$$u^* = \sum_{i=1}^{k} \lambda_i v^i$$

where $k \leq N$ and $v^i \in G$. If it happens that all points in G are extreme points of P, then $k = N$. Therefore

$$J^* = J(u^*) = J\left(\sum_{i=1}^{k} \lambda_i v^i\right)$$

Where $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0$. But J is convex function (given). Hence by definition of convex function

$$J^* = J\left(\sum_{i=1}^{k} \lambda_i v^i\right) \leq \sum_{i=1}^{k} \lambda_i J\left(v^i\right) \tag{1}$$

The above is generalization of $J\left((1-\lambda)u^1 + \lambda u^2\right) \leq (1-\lambda) J\left(u^1\right) + \lambda J\left(u^2\right)$ applied to convex mixtures. Now we look at $J\left(v^i\right)$ term in the above. We pick the maximum of J over all v^i. There must be a point in G where $J(v)$ is largest. We call this value J_G^*. This is the value of

J where it attains its maximum over generator elements $v^i : i = 1 \cdots k$. Eq (1) becomes

$$J^* \leq \sum_{i=1}^{k} \lambda_i J_G^*$$

Where we replaced $J\left(v^i\right)$ by one value, the maximum J_G^*. But J_G^* does not depend on i now, and can take it outside the sum

$$J^* \leq J_G^* \left(\sum_{i=1}^{k} \lambda_i\right)$$

But $\sum_{i=1}^{k} \lambda_i = 1$ by definition. Therefore the above becomes

$$J^* \leq J_G^*$$

We now see that the maximum of $J(u)$ over P is smaller (or equal) than the maximum of $J(u)$ over the generator set G. Hence a maximum occurs at one of the extreme points v^i, since these are by definition taken from G. which is what we are asked to show.

0.5 Problem 5

PROBLEM DESCRIPTION

Barmish

ECE 717 – Homework Optimal Gain

In this homework problem, we consider a modification of the optimal gain scenario defined in class. Now, the performance index includes weighting not only on the state $x(t)$ but also on the on the control $u(t)$. That is, we consider

$$J = \int_0^\infty x^T(t)x(t) + \lambda u^T(t)u(t)dt$$

where $\lambda > 0$ is a given weighting factor.

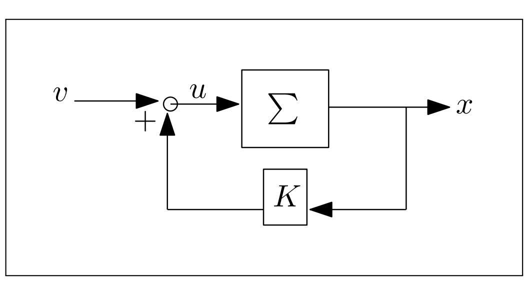
(a) Generalizing upon the approach taken in class, find an expression for the performance $J(K)$ and the associated Lyapunov function which must be satisfied.

(b) Now, using the result from Part (a), we revisit the double integrator problem from class with weighting $\lambda = 1$, initial condition given by $x_1(0) = 1, x_2(0) = 0$ and feedback $K = [k_1\ k_2]$ to be found by optimization. Assuming the two feedback gains are equal (that is, $k_1 = k_2 = k$), find the optimum $k = k^*$, the associated cost J^* and verify that your controller stabilizes the system.

(c) Consider the scenario in Part (b) with the following change: Instead of taking initial condition $x(0)$ as given, assume that each of its components $x_1(0)$ and $x_2(0)$ are independent random variables which are uniformly distributed over $[-1, 1]$. Now find the optimal gain $k = k^*$ minimizing $J(K)$ and the associated optimal cost J^*.

SOLUTION

0.5.1 Part (a)



Let us look at the closed loop. Let $v = 0$ and we have, since $u(t) = kx(t)$

$$\begin{aligned}\dot{x} &= Ax + Bkx\\ &= (A + Bk)\,x\\ &= A_c x\end{aligned}$$

Where A_c is the closed loop system matrix. Since $J(k) = \int_0^\infty x^T(t)\,x(t) + \lambda u^T(t)\,u(t)\,dt$, where $u(t) = kx(t)$, then

$$\begin{aligned}J(k) &= \int_0^\infty x^T x + \lambda (kx)^T (kx)\,dt\\ &= \int_0^\infty x^T x + \lambda x^T \left(k^T k\right) x dt\end{aligned}$$

Let us find a matrix P, if possible such that

$$d\left(x^T P x\right) = -\left(x^T x + \lambda x^T \left(k^T k\right) x\right)$$

Can we find P? Since

$$d\left(x^T P x\right) = x^T P \dot{x} + \dot{x}^T P x$$

Then we need to solve

$$\begin{aligned}x^T P \dot{x} + \dot{x}^T P x &= -\left(x^T x + \lambda x^T \left(k^T k\right) x\right)\\ x^T P (A_c x) + (A_c x)^T P x &= -\left(x^T x + \lambda x^T \left(k^T k\right) x\right)\\ x^T P (A_c x) + \left(x^T A_c^T\right) P x &= -\left(x^T x + \lambda x^T \left(k^T k\right) x\right)\end{aligned}$$

Bring all the x to LHS then

$$\begin{aligned}x^T x + \lambda x^T \left(k^T k\right) x + x^T P (A_c x) + \left(x^T A_c^T\right) P x &= 0\\ \lambda \left(k^T k\right) + P A_c + A_c^T P &= -I\end{aligned}$$

Hence the Lyapunov equation to solve for P is

$$\boxed{\lambda \left(k^T k\right) + P A_c + A_c^T P = -I}$$

Where I is the identity matrix. This is the equation to determine matrix P. Without loss of generality, we insist on P being symmetric matrix. Using this P, now we write

$$\begin{aligned}J(k) &= \int_0^\infty x^T x + \lambda (kx)^T (kx)\,dt\\ &= -\int_0^\infty d\left(x^T P x\right)\\ &= x^T P x\Big|_\infty^0\\ &= x^T(0)\,Px(0) - x^T(\infty)\,Px(\infty)\end{aligned}$$

For stable system, $x(\infty) \to 0$ (since we set $v = 0$, there is no external input, hence if the system is stable, it must end up in zero state eventually). In part (b) we check for k range so that the roots are in the left hand side. Therefore

$$J(k) = x^T(0)\,P(k)\,x(0)$$

With $P(k)$ satisfying solution of Lyapunov equation found above.

0.5.2 Part(b)

For $k = \begin{bmatrix} k_1 & k_2 \end{bmatrix}, x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and system $y'' = u$. Hence $x_1' = x_2, x_2' = u$. Since

$$u = \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The system $\dot{x} = Ax + Bu$ becomes

$$\begin{aligned} x' &= Ax + Bu \\ &= Ax + Bkx \\ &= (A + Bk)\,x \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix} &= \left(\overbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}^{A} + \overbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}^{B} \overbrace{\begin{bmatrix} k_1 & k_2 \end{bmatrix}}^{k} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ k_1 & k_2 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \overbrace{\begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix}}^{A_c} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

For stable system, we need $k_1, k_2 < 0$ from looking at the roots of the characteristic equation. Now we solve the Lyapunov equation.

$$\lambda \left(k^T k \right) + P A_c + A_c^T P = -I$$

$$\lambda \begin{bmatrix} k_1 & k_2 \end{bmatrix}^T \begin{bmatrix} k_1 & k_2 \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix}^T \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\lambda \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \begin{bmatrix} k_1 & k_2 \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix} + \begin{bmatrix} 0 & k_1 \\ 1 & k_2 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\lambda \begin{bmatrix} k_1^2 & k_1 k_2 \\ k_1 k_2 & k_2^2 \end{bmatrix} + \begin{bmatrix} k_1 p_{12} & p_{11} + k_2 p_{12} \\ k_1 p_{22} & p_{21} + k_2 p_{22} \end{bmatrix} + \begin{bmatrix} k_1 p_{21} & k_1 p_{22} \\ p_{11} + k_2 p_{21} & p_{12} + k_2 p_{22} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} k_1 \left(p_{12} + p_{21} + \lambda k_1 \right) & p_{11} + k_1 p_{22} + k_2 p_{12} + \lambda k_1 k_2 \\ p_{11} + k_1 p_{22} + k_2 p_{21} + \lambda k_1 k_2 & \lambda k_2^2 + 2 p_{22} k_2 + p_{12} + p_{21} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

Hence we have 4 equations to solve for $p_{11}, p_{12}, p_{21}, p_{22}$. (but we know also that $p_{12} = p_{21}$). Now let $\lambda = 1$ per the problem, and we obtain the four equations from above as

$$\begin{aligned} k_1^2 + k_1 p_{12} + k_1 p_{21} &= -1 \\ p_{11} + k_1 k_2 + k_1 p_{22} + k_2 p_{12} &= 0 \\ p_{11} + k_1 k_2 + k_1 p_{22} + k_2 p_{21} &= 0 \\ k_2^2 + 2 p_{22} k_2 + p_{12} + p_{21} &= -1 \end{aligned}$$

Solution is (Using Matlab syms).

```
clear;
syms k1 k2 p11 p12 p21 p22;
k   = [k1,k2];
A   = [0,1;0,0];
B   = [0;1];
Ac  = A+B*k;
P   = [p11 p12;p21 p22];
lam = 1;
eq  =  lam*(k.')*k + (Ac.')*P + P*Ac == -eye(2);
sol = solve(eq,{p11,p12,p21,p22});
P   = simplify(subs(P,sol))
x0  = [1;0];
J1  = simplify(x0'*P*x0)
```

.

Gives

```
P =
[ -(k1^3 - k1^2 + k1 - k2^2)/(2*k1*k2),  -(k1^2 + 1)/(2*k1)]
[ -(k1^2 + 1)/(2*k1), -(- k1^2 + k1*k2^2 + k1 - 1)/(2*k1*k2)]
J1 =
-(k1^3 - k1^2 + k1 - k2^2)/(2*k1*k2)
```

$$P=\begin{bmatrix}-\frac{k_1-k_1^2+k_1^3-k_2^2}{2k_1k_2} & -\frac{k_1^2+1}{2k_1}\\ -\frac{k_1^2+1}{2k_1} & -\frac{k_1+k_1k_2^2-k_1^2-1}{2k_1k_2}\end{bmatrix}$$

Hence

$$\begin{aligned}J(k)&=x^T(0)\,P(k)\,x(0)\\&=\begin{bmatrix}1&0\end{bmatrix}\begin{bmatrix}-\frac{k_1-k_1^2+k_1^3-k_2^2}{2k_1k_2} & -\frac{k_1^2+1}{2k_1}\\ -\frac{k_1^2+1}{2k_1} & -\frac{k_1+k_1k_2^2-k_1^2-1}{2k_1k_2}\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}\end{aligned}$$

Therefore

$$\boxed{J(k)=-\frac{1}{2k_1k_2}\left(k_1^3-k_1^2+k_1-k_2^2\right)}$$

For $k_1=k_2=k$, the above becomes

$$\begin{aligned}J(k)&=-\frac{\left(k^3-2k^2+k\right)}{2k^2}\\&=-\frac{\left(k^2-2k+1\right)}{2k}\end{aligned}$$

Or

$$\boxed{J(k)=-\frac{1}{2k}(k-1)^2}$$

Now we find the optimal J^*. Since

$$\frac{dJ(k)}{dk}=\frac{(k-1)^2}{2k^2}-\frac{(2k-2)}{2k}$$

Then $\frac{dJ(k)}{dk}=0$ gives

$$k=1,-1$$

Since k must be negative for stable system, we pick

$$\boxed{k^*=-1}$$

And

$$\frac{d^2J(k)}{dk^2}=\frac{(k-1)^2}{k^3}-\frac{2(1-k)}{k^2}-\frac{1}{k}$$

At $k^*=-1$

$$\frac{d^2J(k)}{dk^2}=1>0$$

Hence this is a minimum. Therefore

$$J^*=-\frac{1}{2k}(k-1)^2\bigg|_{k=-1}$$

Hence

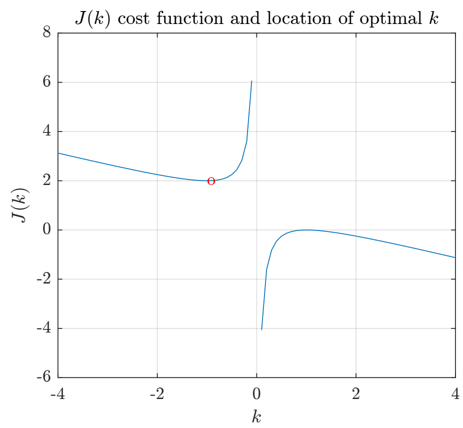
$$\boxed{J^*=2}$$

J^* do not get to zero. (same as in the class problem we did without λu^Tu term. I thought we will get $J^*=0$ now since this I thought it was the reason for using λu^Tu. I hope I did not make mistake, but do not see where if I did. Below is a plot of $J(k)$.

```
clear k;
close all;
reset(0);
set(groot,'defaulttextinterpreter','Latex');
set(groot, 'defaultAxesTickLabelInterpreter','Latex');
set(groot, 'defaultLegendInterpreter','Latex');
f=@(k) (-1./(2*k).*(k-1).^2)
k=-4:.1:4;
plot(k,f(k));
text(-1,f(-1),'o','color','red')
title('$J(k)$ cost function and location of optimal $k$');
```

```
xlabel('$k$'); ylabel('$J(k)$');
grid;
```

.



At $k=1$ then $J(1)=0$, but we can not use $k=1$ since this will make the system not stable. The system now using $k^{*}=-1$ becomes

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

To verify it is stable. Since

$$|(\lambda I - A_c)| = \lambda^2 + \lambda + 1$$

The roots are

$$-\frac{1}{2} \pm \frac{1}{2} i\sqrt{3}$$

Hence the system is stable since real part of roots are negative. If we had used $k=1$, the roots will be $-0.618, 1.618$, and the system would have been unstable.

0.5.3 Part(c)

From last part, we obtained P

$$P = \begin{bmatrix} -\frac{k_1 - k_1^2 + k_1^3 - k_2^2}{2k_1 k_2} & -\frac{k_1^2+1}{2k_1} \\ -\frac{k_1^2+1}{2k_1} & -\frac{k_1 + k_1 k_2^2 - k_1^2 - 1}{2k_1 k_2} \end{bmatrix}$$

When $k_1 = k_2 = k$ the above becomes

$$P = \begin{bmatrix} \frac{-k+2k^2-k^3}{2k^2} & -\frac{k^2+1}{2k} \\ -\frac{k^2+1}{2k} & \frac{1-k-k^3+k^2}{2k^2} \end{bmatrix}$$

Now since $x(0)$ is random variable, then

$$J(k) = E\left(x^T(0)\, P x(0)\right)$$
$$= E\left(\begin{bmatrix} x_1(0) & x_2(0) \end{bmatrix} \begin{bmatrix} \frac{-k+2k^2-k^3}{2k^2} & -\frac{k^2+1}{2k} \\ -\frac{k^2+1}{2k} & \frac{1-k-k^3+k^2}{2k^2} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} \right)$$
$$= E\left(-\frac{1}{2k^2} \left(k^3 x_1^2(0) + 2k^3 x_1(0)\, x_2(0) + k^3 x_2^2(0) - 2k^2 x_1^2(0) - k^2 x_2^2(0) + k x_1^2(0) + 2k x_1(0)\, x_2(0) + k x_2^2(0) - x_2^2(0 \right.\right. \tag{1}$$

Let $E(x_1(0)) = \bar{x}_1$ and $E(x_2(0)) = \bar{x}_2$ Then

$$J(k) = -\frac{1}{2k^2} \left(k^3 \bar{x}_1^2 + 2k^3 \bar{x}_1 \bar{x}_2 + k^3 \bar{x}_2^2 - 2k^2 \bar{x}_1^2 - k^2 \bar{x}_2^2 + k \bar{x}_1^2 + 2k \bar{x}_1 \bar{x}_2 + k \bar{x}_2^2 - \bar{x}_2^2 \right)$$

But $E(x_1(0)) = 0$, hence $\bar{x}_1 = 0$ and similarly $\bar{x}_2 = 0$, but $\bar{x}_1^2 = \frac{1}{3}$ since

$$\int_{-1}^{1} x^2 p(x)\,dx = \frac{1}{2}\int_{-1}^{1} x^2 dx = \frac{1}{2}\left(\frac{x^3}{3}\right)_{-1}^{1} = \frac{1}{3}$$

Similarly $\bar{x}_2^2 = \frac{1}{3}$ and $\bar{x}_1\bar{x}_2 = 0$ (since i.i.d, then $E(x_1(0)\,x_2(0)) = E(x_1(0))\,E(x_2(0)) = 0$. Using these values of expectations, Eq (1) becomes

$$J(k) = -\frac{1}{2k^2}\left(k^3\frac{1}{3} + k^3\frac{1}{3} - 2k^2\frac{1}{3} - k^2\frac{1}{3} + k\frac{1}{3} + k\frac{1}{3} - \frac{1}{3}\right)$$

Or

$$\boxed{J(k) = \frac{-2k^3+3k^2-2k+1}{6k^2}} \tag{2}$$

To find the optimal:

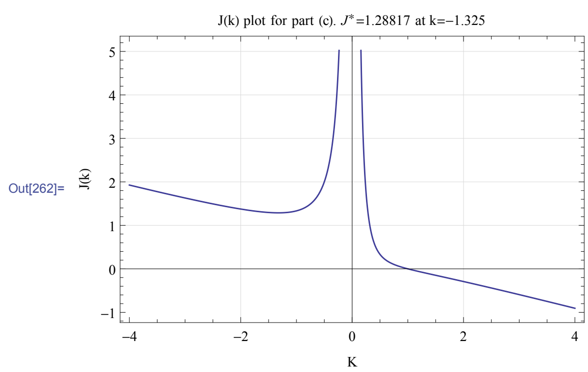
$$\frac{dJ(k)}{dk} = -\frac{1}{3} - \frac{1}{3k^3} + \frac{1}{3k^2}$$

$\frac{dJ(k)}{dk} = 0$ gives 3 roots. The only one which is real and negative (the other two are complex) is

$$\boxed{k^* = -1.325}$$

At this k^*, we check $\frac{d^2J(k)}{dk^2}$ and find it is $0.611 > 0$, hence J is minimum at k^*. The value J^* at k^* is found to be from substituting k^* in (2)

$$\boxed{J^* = 1.28817}$$



We now check if the system is stable. (it should be, since $k^* < 1$). The system now is

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ k_1 & k_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 1 \\ -1.325 & -1.325 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Hence

$$|(\lambda I - A_c)| = \lambda^2 + 1.325\lambda + 1.325$$

The roots are

$$-0.6625 \pm i0.941$$

The system is stable since real part of roots are negative. The following is the step response for system in part(b) and part(c) to compare.

```
%show step responses
close all;
figure();
close all
A   = [0 1;-1 -1];
B   = [1;0]
sys = ss(A,B,[1 0],[0])
step(sys)
hold on;

```

```
A   = [0 1;-1.325 -1.325];
B   = [1;0]
sys = ss(A,B,[1 0],[0])
step(sys)
legend('part(b) step response','part(c) step response')
xlabel('time');
ylabel('y(t)');
grid
```

.