EN.530.603 Applied Optimal Control
HW #1 Solutions

Gowtham Garimella

September 22, 2013

1. To find stationary points and determine whether they are maxima, minima or saddle points: For stationary points, the gradient has to be zero.

   (a)

$$L(x) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$$

$$\nabla L(x) = \begin{bmatrix} \frac{\partial L}{\partial x_1} \\ \frac{\partial L}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2(1 - x_1)(-1) + 200(x_2 - x_1^2)(-2x_1) \\ 2 \times 100(x_2 - x_1^2) \end{bmatrix} = \mathbf{0_{2\times 1}}$$

$$\Rightarrow x_2 = x_1^2 \quad \& \quad -1 + x_1 = 0$$

$$\Rightarrow x_1 = 1 \quad \& \quad x_2 = x_1^2 = 1$$

   Thus the stationary point is (1,1). Now to characterize the stationary point we look at the Hessian of the function:

$$\nabla^2 L(x) = \begin{bmatrix} 2\left[1 - 200x_2 + 600x_1^2\right] & 2(-200x_1) \\ 200(-2x_1) & 200 \end{bmatrix}$$

$$\nabla^2 L(x)|_{(1,1)} = H = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

$$eig(H) = \{1001.6, 0.4\}$$

   Thus the Hessian is positive definite which implies the stationary point (1,1) is a **strict local minima**

   (b)

$$L(u) = (u - 1)(u + 2)(u - 5)$$

$$\nabla L(u) = \frac{\partial L}{\partial u} = (u - 1)(u + 2) + (u - 1)(u - 5) + (u + 2)(u - 5) = 3u^2 - 8u - 7 = 0$$

$$\Rightarrow u^* = \{3.36, -0.69\}$$

Now to characterize the stationary points we look at the Hessian:

$$
\begin{aligned}
\nabla^2 L(u) &= \frac{\partial^2 L}{\partial^2 u} = 6u - 8 \\
\nabla^2 L(u)|_{u=3.36} &= 6(3.36) - 8 = 12.16 \\
\nabla^2 L(u)|_{u=-0.694} &= 6(-0.694) - 8 = -12.16
\end{aligned}
$$

Clearly, point u = 3.36 is **strict local minima** and point u = -0.694 is **strict local maxima**

(c)

$$
\begin{aligned}
L(u) &= (u_1^2 + 3u_1 - 4)(u_2^2 - u_2 + 6) \\
\nabla L(u) = \begin{bmatrix} \frac{\partial L}{\partial u_1} \\ \frac{\partial L}{\partial u_2} \end{bmatrix} &= \begin{bmatrix} (2u_1 + 3)(u_2^2 - u_2 + 6) \\ (u_1^2 + 3u_1 - 4)(2u_2 - 1) \end{bmatrix} = \mathbf{0_{2\times1}} \\
\forall u_2 \in \mathbb{R} \quad & u_2^2 - u_2 + 6 > 0 \\
\Rightarrow u_1 &= -3/2 \\
\Rightarrow u^* &= \begin{bmatrix} -3/2 \\ 1/2 \end{bmatrix}
\end{aligned}
$$

Now that we found the stationary point let us characterize it by looking at the Hessian of the above objective function:

$$
\begin{aligned}
\nabla^2 L(u) &= \begin{bmatrix} 2(u_2^2 - u_2 + 6) & (2u_1 + 3)(2u_2 - 1) \\ (2u_1 + 3)(2u_2 - 1) & 2(u_1^2 + 3u_1 - 4) \end{bmatrix} \\
\nabla^2 L(u)|_{u=(-1.5,0.5)} &= \begin{bmatrix} 11.5 & 0 \\ 0 & -12.5 \end{bmatrix} \\
\Rightarrow eig(\nabla^2 L(u)) &= \{11.5, -12.5\}
\end{aligned}
$$

Since the eigen values are both positive and negative at the stationary point u = (-3/2, 1/2), the point is a **saddle point**.

2. To find the stationary points and determine the maxima, minima or saddle points:

(a) Define L' using lagrangian multipliers as $L' := L(x) + \lambda f(x)$ where $\lambda \in \mathbb{R}$ is a scalar. To find the stationary points, we have to equate the gradient to zero:

$$L'(x) = \frac{1}{2}(x_1^2 + x_2^2 + x_3^2) + \lambda(x_1 + x_2 + x_3)$$
$$\frac{\partial L'}{\partial x_1} = x_1 + \lambda = 0 \quad \text{(at stationary point)}$$
$$\Rightarrow \lambda = -x_1^* \quad (x_1^* \text{ is the stationary point})$$
$$\frac{\partial L'}{\partial x_2} = x_2 + \lambda = 0$$
$$\frac{\partial L'}{\partial x_3} = x_3 + \lambda = 0$$

From constraint f(x)

$$\Rightarrow x_2^* = x_3^* = x_1^* = \lambda = 0$$

Thus the stationary point is (0,0,0). To characterize the stationary point we look at the Hessian of the L'. Since $\lambda = 0$, $L' = L$. Thus

$$\nabla^2 L' = \nabla^2 L = H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow eig(H) = \{1, 1, 1\}$$

Since the eigen values are all positive, the Hessian is positive definite and hence the stationary point (0,0,0) is a **strict local minima**

(b) As in above equation define the augmented cost function and let's equate the gradient to zero:

$$L'(u) = L(u) + \lambda f(u) = (u_1^2 + 3u_1 - 4)(u_2^2 - u_2 + 6) + \lambda(u_1 - 2u_2)$$
$$\frac{\partial L'}{\partial u_1} = (2u_1 + 3)(u_2^2 - u_2 + 6) + \lambda = 0 \quad \text{(at stationary point)}$$
$$\frac{\partial L'}{\partial u_2} = (u_1^2 + 3u_1 - 4)(2u_2 - 1) - 2\lambda = 0 \quad \text{(at stationary point)}$$
$$\frac{\partial L'}{\partial \lambda} = u_1 - 2u_2 = 0$$
$$\Rightarrow u_1 = 2u_2 \quad \text{and}$$
$$\Rightarrow (u_1^2 + 3u_1 - 4)(2u_2 - 1) + 2(2u_1 + 3)(u_2^2 - u_2 + 6) = 0$$
$$\Rightarrow (2u_2^2 + 3u_2 - 2)(2u_2 - 1) + (4u_2 + 3)(u_2^2 - u_2 + 6) = 0$$
$$\Rightarrow 8u_2^3 + 3u_2^2 + 14u_2 + 20 = 0$$

The above cubic equation yields only one real root i.e $u_2 = -1.03$ i.e $u = (-0.515, -1.03)$. To characterize this stationary point we look at the Hessian of L':

$$\nabla^2 L' = \begin{bmatrix} 2(u_2^2 - u_2 + 6) & (2u_1 + 3)(2u_2 - 1) \\ (2u_1 + 3)(2u_2 - 1) & 2(u_1^2 + 3u_1 - 4) \end{bmatrix}$$
$$\nabla^2 L'|_{u=(-0.515,-1.03)} = H = \begin{bmatrix} 16.182 & -6.028 \\ -6.028 & -10.56 \end{bmatrix}$$
$$eig(H) = \{17.48, -11.86\}$$

Since the eigen values of the Hessian are positive and negative, the stationary point(-0.515,-1.03) is a **saddle point**.

3. To optimize the quadratic cost function with constraints given as:

$$L(x,u) = \frac{1}{2}x^T Q x + \frac{1}{2}u^T R u$$
$$f(x,u) = Ax + Bu + c = 0 \quad x \in \mathbb{R}^n, c \in \mathbb{R}^m$$

Given Q and R are positive definite matrices. Define symmetric positive definite matrices $\bar{Q}, \bar{R}$ s.t

$$\bar{Q} = \frac{1}{2}(Q + Q^T); \quad \bar{R} = \frac{1}{2}R + R^T$$

It is obvious to check that the cost function can be written as:

$$L(x,u) = \frac{1}{2}x^T \bar{Q} x + \frac{1}{2}u^T \bar{R} u$$

Using the Lagrangian multipliers method:

$$L'(x,u) = \frac{1}{2}x^T \bar{Q} x + \frac{1}{2}u^T \bar{R} u + \lambda^T(Ax + Bu + c)$$
$$\frac{\partial L'}{\partial x} = x^T \bar{Q} + \lambda^T(A) = \mathbf{0_{1\times n}}$$
$$\frac{\partial L'}{\partial u} = u^T \bar{R} + \lambda^T(B) = \mathbf{0_{1\times m}}$$
$$\frac{\partial L'}{\partial \lambda} = Ax + Bu + c = 0$$

Assuming A is invertible,

$$\Rightarrow \lambda^T = -x^T \bar{Q} A^{-1}$$
$$\Rightarrow u^T \bar{R} - x^T \bar{Q} A^{-1} B = \mathbf{0_{1\times m}}$$
$$\Rightarrow -B^T {A^T}^{-1} \bar{Q} x + \bar{R} u = 0 \quad \text{(Since } \bar{Q} \text{ and } \bar{R} \text{ are symmetric)}$$
$$\text{and } Ax + Bu + c = 0$$

Thus the necessary first order condition for minimization assuming A is invertible is:

$$\begin{bmatrix} -B^T {A^T}^{-1} \bar{Q} & \bar{R} \\ A & B \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ -c \end{bmatrix}$$

Second order sufficient condition requires that $\nabla^2 L'$ is positive definite i.e:

$$\nabla^2 L' = H = \begin{bmatrix} \bar{Q} & \mathbf{0_{n\times m}} \\ \mathbf{0_{m\times n}} & \bar{R} \end{bmatrix}$$

Since $\bar{Q}$ and $\bar{R}$ are both positive definite,

$$\begin{bmatrix} x & u \end{bmatrix} \begin{bmatrix} \bar{Q} & \mathbf{0_{n\times m}} \\ \mathbf{0_{m\times n}} & \bar{R} \end{bmatrix} \begin{bmatrix} x^T \\ u^T \end{bmatrix} = x^T\bar{Q}x + u^T\bar{R}u > 0$$

Thus the Hessian is already positive definite and hence there is no required second order sufficiency condition. Which implies that the necessary and sufficient condition for minimizing the function assuming A is invertible is:

$$\begin{bmatrix} -B^T A^{T^{-1}}\bar{Q} & \bar{R} \\ A & B \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ -c \end{bmatrix}$$

4. Minimization problem stated as:

$$\begin{aligned} f^k(d) &= f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d \quad d \in \mathbb{R}^n \\ s.t \quad \|d\| &\leq \gamma^k \end{aligned}$$

If the constraint is active it is treated as an equality and following Lagrange multipliers method:

$$\begin{aligned} f(d) &= \|d\| - \gamma^k = \sqrt{d^T d} - \gamma^k \\ L'(d) &= f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d + \lambda(\sqrt{d^T d} - \gamma^k) \quad \lambda \geq 0 \\ \frac{\partial L'}{\partial d} &= \nabla f(x^k)^T + d^T \nabla^2 f(x^k) + \lambda\left(\frac{1}{2\sqrt{d^T d}} d^T\right) = 0 \\ \frac{\partial L'}{\partial \lambda} &= \sqrt{d^T d} - \gamma^k = 0 \\ \Rightarrow -\nabla f(x^k)^T &= d^T \left(\nabla^2 f(x^k) + \frac{\lambda}{2\gamma^k} I \right) \\ \Rightarrow \left(\nabla^2 f(x^k) + \delta^k I\right) d &= -\nabla f(x^k) \quad \text{Since hessian and Identity are both symmetric} \end{aligned}$$

Thus we have shown that the constrained optimization problem is equivalent to solving the above form of matrix equations for δ^k.

The Hessian for the augmented cost function is $\nabla^2 f(x^k) + \delta^k I$. Thus we need to choose the right δ^k so that the Hessian is **positive definite**. A reasonable choice of δ^k is $max\{-eig(\nabla^2 f(x^k))\}$. Since the eigen values of the augmented cost function is the sum of original eigenvalues and δ^k. Such a choice will ensure all the eigen values of Hessian will be positive and hence positive definite

Acknowledgements

I hereby declare that I have not discussed this homework with anyone. The solutions written here are my own work and from lecture notes and sample code provided by the professor.

Gowtham Garimella