1. Find the stationary points (i.e. that satisfy $\nabla L = 0$) of the following and determine whether they are maxima, minima, or saddle points:

   (a) $L(x) = (1 - x_1)^2 + 200(x_2 - x_1^2)^2$
   (b) $L(u) = (u - 1)(u + 2)(u - 3)$
   (c) $L(u) = (u_1^2 + 3u_1 - 4)(u_2^2 - u_2 + 3)$

2. Find the stationary points of the following and determine whether they are maxima, minima, or saddle points:

   (a) 
   \[
   \begin{align*}
   &\text{minimize } L(x) = \frac{1}{2} \left( x_1^2 + x_2^2 + x_3^2 \right) \\
   &\text{subject to } f(x) = x_1 + x_2 + x_3 = 0
   \end{align*}
   \]

   (b) 
   \[
   \begin{align*}
   &\text{minimize } L(u) = \left( u_1^2 + 3u_1 - 4 \right)(u_2^2 - u_2 + 3) \\
   &\text{subject to } f(u) = u_1 - 2u_2 = 0
   \end{align*}
   \]

3. (a) Consider the optimization of a quadratic cost subject to linear constraints, i.e. minimize 

   \[
   L(x,u) = \frac{1}{2} x^T Qx + \frac{1}{2} u^T Ru + s^T x,
   \]

   subject to 

   \[
   f(x,u) = Ax + Bu + c = 0,
   \]

   where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$; $Q \geq 0$ (positive semidefinite matrix) and $R > 0$ (positive definite matrix); $A \in \mathbb{R}^{nxn}$, $B^{nxm}$, $s \in \mathbb{R}^n$ and $c \in \mathbb{R}^n$.

   • Derive the necessary and sufficient conditions for an optimal solution using the Lagrangian multiplier approach. Be careful which matrices you are allowed to invert.

   • Assume that $A$ is full rank and compute the actual optimal solution.
(b) Consider the optimization of a quadratic cost subject to linear constraints, i.e. minimize

\[ L(y) = \frac{1}{2} y^T M y + k^T y \]

subject to

\[ f(y) = Ay + c = 0, \]

where \( y \in \mathbb{R}^n \), \( M > 0 \) is positive definite, \( A \in \mathbb{R}^{m \times n} \) for \( m < n \) is full rank, \( k \in \mathbb{R}^n \), and \( c \in \mathbb{R}^m \). Compute the optimal solution \( y^* \) and show that it is a global minimum.

4. Implementation: you are free to use parts of the code provided at the course homepage.

(a) Write a MATLAB function which implements gradient descent to optimize the cost-function in problem 1a given by

\[ L(x) = (1 - x_1)^2 + 200(x_2 - x_1^2)^2. \]

You can use either a constant or a variable stepsize. What is the effect of the step-size choice? Use a starting point at \( x = (0, 0) \).

(b) Write a MATLAB function which implements Newton’s method to optimize the function in problem 1a. Comment on the rate of convergence of Newton’s method compared to gradient descent.

(c) Optimize the cost-function

\[ L(x) = x_1 e^{-x_1^2 - \frac{1}{2} x_2^2} + x_1^2/10 + x_2^2/10 \]

using gradient descent and Newton’s method using the starting point \( x = (1.5, 3.0) \). Compare the performance of the two algorithms.

(d) Find analytically the optimum of the following problem:

\[
\begin{align*}
\text{minimize} & \quad L(x, u) = x^2 + 4x + 8u^2; \\
\text{subject to} & \quad f(x, u) = x - 7u + 5 \leq 0
\end{align*}
\]

and use MATLAB fmincon function to verify your solution.

Note: email your code to spatkar1@jhu.edu and marin@jhu.edu with subject line starting with: EN530.603.F2017.HW1; in addition attach a printout of the code to your homework solutions.

5. Consider the minimization of \( f(x) \) for \( x \in \mathbb{R}^n \). Newton’s method is derived by finding the direction \( d^k \in \mathbb{R}^n \) which minimizes the local quadratic approximation \( f^k \) of \( f \) at \( x^k \) defined by

\[ f^k(d) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2 f(x^k) d. \]

In contrast, the search direction \( d^k \) in a trust-region Newton method is derived by solving the constrained optimization

\[ \text{minimize } f^k(d) \text{ subject to } \| d \| \leq \gamma^k, \]
for a given $\gamma^k > 0$ called the trust-region radius. Using the Lagrangian multiplier approach prove that this optimization is equivalent to solving

$$(\nabla^2 f(x^k) + \delta^k I) d^k = -\nabla f(x^k),$$

where $\delta^k \geq 0$. How do you interpret the value $\delta^k$. Can you propose a reasonable choice for $\delta^k$ considering the properties of $\nabla^2 f(x^k)$.

6. Read one (or both) of the following two historical perspectives on optimal control and write a one-paragraph summary of the paper you choose:


(b) Sussman and Willems, “300 Years of Optimal Control: From the Brachystochrone to the Maximum Principle”, IEEE Control Systems, 1997