Solutions to Homework 1

February 24, 2010

1. **Bertsekas 1.1.1:** Given the function

$$f(x, y) = x^2 + y^2 + \beta xy + x + 2y,$$

we need to determine which of its stationary points are minima. From the set of necessary conditions $\nabla f = 0$, we get

$$\begin{aligned} 2x + \beta y &= -1 \\ \beta x + 2y &= -2. \end{aligned}$$

When these simultaneous equations are solved for (x, y), we obtain

$$x = \frac{2(1-\beta)}{(\beta+2)(\beta-2)},\ y = \frac{4-\beta}{(\beta+2)(\beta-2)}.$$

The sufficient conditions for a minimum require that the Hessian $\nabla^2 f$ be positive definite. From this condition, we obtain

$$|\beta| < 2.$$

When $|\beta| = 2$, there are no available solutions for the pair (x^*, y^*).

2. **Bertsekas 1.1.2:**

(a) The function $f(x, y) = (x^2 - 4)^2 + y^2$ has stationary points $(0, 0)$, (2,0) and (-2,0). The point (0,0) is a saddle point since $\frac{\partial^2 f}{\partial x^2} < 0$ and $\frac{\partial^2 f}{\partial y^2} > 0$.

(b) The function $f(x, y) = \frac{1}{2}x^2 + x\cos y$ has stationary points $\{(0, (2k+1)\frac{\pi}{2}),\ k = 0, \pm 1, \pm 2 \ldots\}$, $\{(-1, 2k\pi),\ k = 0, \pm 1, \pm 2, \ldots\}$ and $\{(1, (2k+1)\pi), k = 0, \pm 1, \pm 2, \ldots\}$. The sufficient condition for a minimum requires that

$$-x\cos^2 y - \sin^2 y \geq 0$$

which is only satisfied for the sets $\{(-1, 2k\pi),\ k = 0, \pm 1, \pm 2, \ldots\}$ and $\{(1, (2k+1)\pi), k = 0, \pm 1, \pm 2, \ldots\}$.

(c) Given the function $f(x, y) = \sin(x) + \sin(y) + \sin(x + y)$, the set of stationary points within $\{(x, y)|0 < x < 2\pi, 0 < y < 2\pi\}$ are given by the equations

$$\begin{aligned} \cos x + \cos(x + y) &= 0, \\ \cos y + \cos(x + y) &= 0. \end{aligned}$$

The solutions within $0 < x < 2\pi$ and $0 < y < 2\pi$ are (π, π), $(\pi/3, \pi/3)$, $(5\pi/3, 5\pi/3)$. The matrix of second derivatives is

$$\begin{bmatrix} -\sin x - \sin(x+y) & -\sin(x+y) \\ -\sin(x+y) & -\sin y - \sin(x+y) \end{bmatrix}.$$

We can easily check that $(5\pi/3, 5\pi/3)$ is the only local minimum and that $(\pi/3, \pi/3)$ is the only local maximum. (See attached MATLAB plot).

(d) The stationary points of the function $f(x, y) = (y - x^2)^2 - x^2$ are $(0, 0)$. The Hessian at this stationary point is $\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$ which implies that the point $(0, 0)$ is a saddle point.

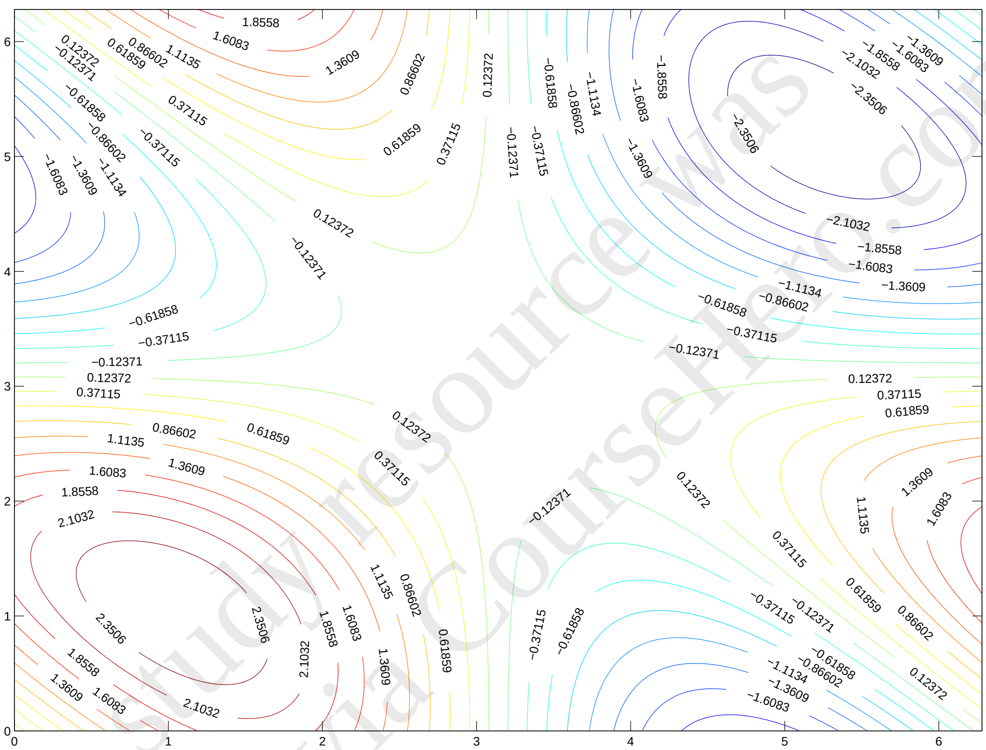

Figure 1: MATLAB plot of the level sets of the function $f(x, y) = \sin x + \sin y + \sin(x + y)$.

(e) Since the unconstrained problem has no local minima, the constrained problem must have minima on the boundaries $y = \pm 1$. Performing 1D calculus on the functions $f(x) = (1-x^2)^2 - x^2$ and $g(x) = (1+x^2)^2 - x^2$, we see that the minima occur at $(0, -1)$, $(\sqrt{6}/2, 1)$ and $(-\sqrt{6}/2, 1)$.

3. **Bertsekas 1.1.3:**

(a) Since the function $f(x^* + \alpha d)$ is minimized at $\alpha = 0$, we must have as a necessary condition that the derivative w.r.t. α is zero. This gives us $\nabla f^T d = 0$ and since this has to be zero for all d, we get $\nabla f = 0$.

(b) Consider the function $f(y, z) = (z - py^2)(z - qy^2)$ where $0 < p < q$. For an $x^* = (0, 0)$, $x^* + \alpha d = (\alpha d_1, \alpha d_2)$. Then

$$f(\alpha d_1, \alpha d_2) = \alpha^2 d_2^2 - \alpha^3 d_2 (p+q) d_1^2 + pq\alpha^4 d_1^2 d_2^2.$$

From $\frac{\partial f}{\partial \alpha} = 0$, we get $\alpha = 0$ to be a stationary point. From the condition $\frac{\partial^2 f}{\partial \alpha^2} \geq 0$, we get $2d_2^2 \geq 0$ which is strictly greater than zero for any $d_2 \neq 0$. Furthermore $f(y, my^2) = (m-p)(m-q)y^4 < 0$ if $y \neq 0$ and $p < m < q$. Since y can be an arbitrarily small ϵ, we get $f(\epsilon, m\epsilon^2) = (m-p)(m-q)\epsilon^4 < 0$ despite the fact that $(0, 0)$ is a local minimum along *every line* passing through the origin. This illustrates that minimization in a subspace—every line in this case—is not sufficient to be a local minimum in the original space.

4. **Bertsekas 1.1.6:** We seek to minimize

$$f(x) = \sum_{i=1}^{m} w_i \|x - y_i\|$$

subject to $x \in \mathbb{R}^n$ and where $w_1, \ldots, w_m$ are given positive scalars.

(a) First we show that the objective function is convex. We require $f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$.

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &= \sum_{i=1}^{m} w_i \|\alpha x_1 + (1-\alpha)x_2 - y_i\| \\ &\leq \sum_{i=1}^{m} \alpha w_i \|x_1 - y_i\| + \sum_{i=1}^{m} (1-\alpha) w_i \|x_2 - y_i\| \\ &\leq \alpha f(x_1) + (1-\alpha) f(x_2). \end{aligned}$$

We have used the triangle inequality in step 2. By Proposition 1.1.2, a local minimum is also a global minimum.

(b) Since $f(x)$ may not be strictly convex, the global minimum may not be unique.

(c) Assume that the lengths of the m strings are the same (d). Then the height of each string is $h_i = h - (d - \|x - y_i\|)$ where h is the height of the table. Consequently

$$\begin{aligned} f(x) &= \sum_{i=1}^{m} w_i \|x - y_i\| \\ &= \sum_{i=1}^{m} w_i (h_i - h - d) \\ &\propto \sum_{i=1}^{m} w_i h_i. \end{aligned}$$

Therefore the minimization of $f(x)$ is equivalent to the minimization of the potential energy $\sum_{i=1}^{m} w_i h_i$.

5. **Bertsekas 1.1.10:** The function

$$f(x) = x_2^2 - a x_2 \|x\|^2 + \|x\|^4 = \left(\|x\|^2 - \frac{a}{2} x_2\right)^2 + \left(1 - \frac{a^2}{4}\right) > 0$$

for $0 < a < 2$. For the points $(-p, q)$ and (p, q), we have that

$$f(x) = q^2 - aq(p^2 + q^2) + (p^2 + q^2)^2.$$

We need to show that $f(p,q) = f(-p,q) < f(0,q)$. From direct algebra,

$$f(p,q) - f(0,q) = p^2 + 2(q - \frac{a}{4})^2 - \frac{a^2}{8} \leq 0$$

for $q = \frac{a}{4}$ and $0 < p \leq \frac{a}{2\sqrt{2}}$. For $p = \frac{a}{2\sqrt{2}}$ and $q = \frac{a}{4}$, we get $f(p,q) = \frac{a^2}{16} - \frac{3a^4}{256} = \frac{a^2}{16}\left(1 - \frac{3a^2}{16}\right) = \gamma$. For $0 < a < 2$, $\gamma > 0$.

6. **Bertsekas 1.2.7:** The engineer's approach basically amounts to a coordinate ascent strategy. She performs local maximization w.r.t. one variable while holding the other fixed and vice versa. Local maximization w.r.t. one variable is similar to steepest ascent except that the ascent is not clocked: each variable is separately optimized while holding the other fixed. As long as each step is guaranteed to increase or keep same the current I, this process will perform a type of ascent on the current. However, one cannot usually guarantee convergence of each variable to a fixed point.