ECE 6553: Homework #1

Klaus Okkelberg

January 26, 2017

1. (a) Consider vectors $\epsilon v \in \mathbb{R}^m$, $\epsilon \in \mathbb{R}$, $\|v\| = 1$. Notice that ϵ is the magnitude and v is the direction of the vector. The size of the growth of $g(u)$ in the direction of ϵv is given by

$$\begin{aligned} g(u+\epsilon v) - g(u) &= g(u) + \epsilon \frac{\partial g}{\partial u}(u) v + o(\epsilon) - g(u) \\ &= \epsilon \frac{\partial g}{\partial u}(u) v + o(\epsilon) \\ &= \epsilon \frac{\partial g^{\mathrm{T}}}{\partial u}(u) \cdot v + o(\epsilon) \\ &= \epsilon \nabla g(u) \cdot v + o(\epsilon) \\ &= \epsilon \|\nabla g(u)\| \, \|v\| \cos\theta + o(\epsilon) \\ &= \epsilon \|\nabla g(u)\| \cos\theta + o(\epsilon), \end{aligned}$$

where θ is the angle between $\nabla g(u)$ and v. We can see that the amount of increase is maximized when $\theta = 0$, i.e. v points along $\nabla g(u)$. Therefore, g grows the most in the direction of $\nabla g(u)$. $\square$

(b) Let $u = r(t)$, $t \in \mathbb{R}$, be the curve that satisfies the constraint $h(u) = 0$, i.e. $h(r(t)) = 0$ $\forall t$. Taking the derivative of the constraint using the chain rule,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} h(r(t)) &= \frac{\partial h(r(t))}{\partial r(t)} \frac{\mathrm{d}r(t)}{\mathrm{d}t} \\ &= \frac{\partial h(u)}{\partial u} r'(t) \\ &= \nabla h(u) \cdot r'(t) = 0 \end{aligned}$$

Notice that the derivative of the curve $r'(t)$ is the tangent plane to the constraint set at u. Since the dot product of the gradient and $r'(t)$ is 0, the gradient $\nabla g(u)$ is orthogonal to the tangent plane to the constraint set at u. $\square$

2. The Lagrangian is

$$L = (u_1 - 2)^2 + 2(u_2 - 1)^2 + \lambda_1(u_1 + 4u_2 - 3) + \lambda_2(u_2 - u_1)$$

The FONCs are

$$\begin{aligned} \frac{\partial L}{\partial u_1} &= 2(u_1 - 2) + \lambda_1 - \lambda_2 = 0 \\ \frac{\partial L}{\partial u_2} &= 4(u_2 - 1) + 4\lambda_1 + \lambda_2 = 0 \\ \lambda_1(u_1 + 4u_2 - 3) &= 0 \\ \lambda_2(u_2 - u_1) &= 0 \\ u_1 + 4u_2 - 3 &\le 0 \\ u_2 - u_1 &\le 0 \\ \lambda_1 &\ge 0 \\ \lambda_2 &\ge 0 \end{aligned}$$

Consider the following combinations in which the constraints can be active/inactive:

(a) $\lambda_1 = \lambda_2 = 0$

$$\begin{cases} 2(u_1 - 2) = 0 \\ 4(u_2 - 1) = 0 \end{cases} \implies \begin{cases} u_1 = 2 \\ u_2 = 1 \end{cases}$$

Since $u_1 + 4u_2 \nleq 3$, this solution is not feasible.

(b) $u_1 + 4u_2 - 3 = 0$, $\lambda_2 = 0$

$$\begin{cases} 2(u_1 - 2) + \lambda_1 = 0 \\ 4(u_2 - 1) + 4\lambda_1 = 0 \\ u_1 + 4u_2 = 3 \end{cases} \implies \begin{bmatrix} u_1 \\ u_2 \\ \lambda_1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 4 & 4 \\ 1 & 4 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 5/3 \\ 1/3 \\ 1 \end{bmatrix}$$

We can verify that the constraints hold:

$$\begin{aligned} u_1 + 4u_2 &\leq 3 \\ u_1 &\geq u_2 \end{aligned}$$

Therefore, $u_1 = 5/3$, $u_2 = 1/3$ is a local minimizer.

(c) $\lambda_1 = 0$, $u_1 = u_2$

$$\begin{cases} 2(u_1 - 2) - \lambda_2 = 0 \\ 4(u_2 - 1) + \lambda_2 = 0 \\ u_1 = u_2 \end{cases} \implies \begin{cases} u_1 = u_2 = 4/3 \\ \lambda_2 = -4/3 \end{cases}$$

Since $\lambda_2 < 0$, this solution is not feasible.

(d) $u_1 + 4u_2 - 3 = 0$, $u_1 = u_2$

$$\begin{cases} 2(u_1 - 2) + \lambda_1 - \lambda_2 = 0 \\ 4(u_2 - 1) + 4\lambda_1 + \lambda_2 = 0 \\ u_1 + 4u_2 = 3 \\ u_1 = u_2 \end{cases} \implies \begin{cases} u_1 = u_2 = 3/5 \\ \lambda_1 = \\ \lambda_2 = \end{cases}$$

this solution is not feasible.

Since the optimization problem only has one feasible solution, the global minimum is given by

$$\begin{aligned} u_1^* &= 5/3 \\ u_2^* &= 1/3 \end{aligned}$$

3. (a) The minimization problem is

$$\begin{aligned} &\min \alpha u_1^2 + \beta u_2^2 \\ &\text{s.t. } u_1 + u_2 = q \end{aligned}$$

The Lagrangian is

$$L = \alpha u_1^2 + \beta u_2^2 + \lambda(u_1 + u_2 - q)$$

The FONCs are

$$\begin{aligned}
\frac{\partial L}{\partial u_1} &= 2\alpha u_1 + \lambda = 0 \\
\frac{\partial L}{\partial u_2} &= 2\beta u_2 + \lambda = 0 \\
\frac{\partial L}{\partial \lambda} &= u_1 + u_2 - q = 0
\end{aligned}$$

From the first two conditions, $\alpha u_1 = \beta u_2$. Substituting this into the third condition gives

$$\begin{aligned}
u_1 &= \frac{\beta}{\alpha+\beta} q \\
u_2 &= \frac{\alpha}{\alpha+\beta} q
\end{aligned}$$

(b) It is clear that the highest cost is achieved by only shipping through the more expensive option, i.e.

$$\begin{cases} \max \alpha u_1^2 + \beta u_2^2 \\ \text{s.t. } u_1 + u_2 = q \end{cases} \implies \begin{cases} u_i^* = q, & i = \arg\max(\alpha,\beta) \\ u_j^* = 0, & j = \arg\min(\alpha,\beta) \end{cases}$$

This gives a total cost of $\max(\alpha,\beta)q^2$. The cost of the minimizers from the previous part is

$$\alpha \left(\frac{\beta}{\alpha+\beta} q\right)^2 + \beta \left(\frac{\alpha}{\alpha+\beta} q\right)^2 = \frac{\alpha\beta}{\alpha+\beta} q^2 < \min(\alpha,\beta)q^2 \le \max(\alpha,\beta)q^2$$

Therefore, the combination found in the previous part is not the worst combination and must be the best combination instead.

4. By convexity of g, for all $u \in \mathbb{R}^m$ and $\alpha \in [0,1]$,

$$\begin{aligned}
g(\alpha u + (1-\alpha)u^*) &\le \alpha g(u) + (1-\alpha) g(u^*) \\
g(\alpha u + (1-\alpha)u^*) &\le g(u^*) + \alpha(g(u) - g(u^*)) \\
g(u) - g(u^*) &\ge \frac{g(\alpha u + (1-\alpha)u^*) - g(u^*)}{\alpha}
\end{aligned}$$

Taking the limit as $\alpha \to 0$,

$$\begin{aligned}
g(u) - g(u^*) &\ge \lim_{\alpha\to 0} \frac{g(\alpha u + (1-\alpha)u^*) - g(u^*)}{\alpha} \\
&= \left[\frac{\mathrm{d}}{\mathrm{d}\alpha}\Big\{g(\alpha u + (1-\alpha)u^*) - g(u^*)\Big\}\right]_{\alpha=0} \\
&= \frac{\partial g}{\partial u}(u^*) \left[\frac{\mathrm{d}}{\mathrm{d}\alpha}(\alpha u + (1-\alpha)u^*)\right]_{\alpha=0} \\
&= \frac{\partial g}{\partial u}(u^*)(u - u^*) = 0 \\
g(u) &\ge g(u^*),
\end{aligned}$$

where L'Hôpital's rule was used to calculate the limit. Since $g(u^*) \leq g(u)\ \forall u \in \mathbb{R}^m$, u^* is the global minimum to g.

5.

6.