Francisco Javier Sáez Maldonado
77448344F
franciscojavier.saez@estudiante.uam.es

Take Home Exam 2
Convex Optimization

2022-03-25

Problem 1
We have worked out the elementary vision of Lagrange multipliers, assuming that , from $g(x, y) = 0$, we can find a function $y = h(x)$ such that $g(x, h(x)) = 0$.

But sometimes, what we get is that there is an h such that $g(x, h(x)) = 0$. Rewrite the Lagrange multiplier analysis in the lecture slides under this assumption.

Consider $f, g : \mathbb{R}^2 \to \mathbb{R}$, and the following minimization problem:

$$\min f(x, y) \;\; \text{s.t. } g(x, y) = 0 \tag{1}$$

Now, we can use the **implicit function theorem** to find a dependence between the variables of the restriction. This theorem (not completely formally) **states** the following: let $g : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a continuously differentiable function, $(x, y) \in \mathbb{R}^{n+m}$ such that $g(x, y) = 0$. If the jacobian with respect to the variables in y is invertible, then there exists an open subset U such that $h(x) = y$ and $g(x, h(x)) = 0$ for all $x \in U$.

Assuming that the **conditions for this theorem are matched**, we can apply it to the **jacobian with respect to the variables in y** to obtain an U' where $h(y) = x$ and $f(h(y), y) = 0$ for all $y \in U$. Thus, we can write:

$$f(x, y) = f(h(y), y) = \psi(y)$$

The, we can keep the procedure as it is done in the slides. Let us see this:

Consider that y^* is a minimum with $x^* = h(y^*)$. Then, we have:

$$0 = \psi'(y^*) = \frac{\partial f}{\partial x}(x^*, y^*)h'(y^*) + \frac{\partial f}{\partial y}(x^*, y^*).$$

Using that (x^*, y^*) is a minimum and that $g(h(y), y) = 0$, we have that:

$$0 = \frac{\partial g}{\partial x}(x^*, y^*)h'(y^*) + \frac{\partial g}{\partial y}(x^*, y^*) \implies h'(y^*) = -\frac{\frac{\partial g}{\partial y}(x^*, y^*)}{\frac{\partial g}{\partial x}(x^*, y^*)}$$

Problem 2
We want to solve the following constrained restriction problem:

$$\begin{aligned} \min \quad & x^2 + 2xy + 2y^2 - 3x + y \\ \text{s.t} \quad & x + y = 1 \\ & x, y \geq 0. \end{aligned}$$

Argue first that f is convex and then:

1. Write its Lagrangian with α, β the multipliers of the inequality constraints.
2. Write the KKT conditions.
3. Use them to solve the problem. For this consider separately the $(\alpha = \beta = 0)$, $(\alpha > 0, \beta = 0)$, $(\alpha = 0, \beta > 0)$, $(\alpha > 0, \beta > 0)$ cases.

Let us first see that f is convex. We know that a characterization of convex functions is that they have a definite positive hessian matrix Hf. Firstly, we observe that the gradient of f is

$$\nabla f(x, y) = (2x + 2y - 3, 2x + 4y + 1)\,.$$

Hence, the hessian of f is

$$Hf = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$$

This matrix is positive definite, since it is Hermitian (it is real-valued and symmetric) and we can apply the **Sylvester's criterion** to see that all the minor determinants (which are 2 and 4) are positive. Hence, f is convex. Now, given that f is convex, we can make use of the **theorem** that states that if $f, g_i \in C^1$ and convex, and h_j are affine, then a KKT point x^* is an optimum of the original problem. Since our problem matches these conditions, we will find the candidates to KKT points.

The **Lagrangian** of f in terms of α, β, λ is:

$$L(x, y, \alpha, \beta, \lambda) = x^2 + 2xy + 2y^2 - 3x + y + \lambda(x + y - 1) + \alpha x + \beta y.$$

Now, to write the KKT conditions. As a very brief summarization, the KKT conditions are:

$$\begin{aligned} \nabla f(x^*) + \sum_i^m \lambda_i \nabla g_i(x^*) + \sum_i^p \mu_j \nabla h_j(x^*) = 0, \\ \lambda_i g_i(x^*) = 0 \end{aligned}$$

With $\lambda_i \geq 0$. If we apply them to our problem, **our KKT conditions** are:

$$\begin{aligned} 2x + 2y - 3 + \lambda + \alpha = 0 \\ 2x + 4y + 1 + \lambda + \beta = 0 \\ \alpha x = 0 \\ \beta y = 0 \end{aligned}$$

Now, we want to solve this equations system to see if we can find a minimum of our problem. We have the following cases:

- Case $\alpha = \beta = 0$.
  In this case, the system is

  $$\begin{aligned} 2x + 2y - 3 + \lambda = 0 \\ 2x + 4y + 1 + \lambda = 0 \end{aligned}$$

  If we substitute $2x$ from the first equation into the second one, we obtain

  $$-2y + 3 - \lambda + 4y + 1 + \lambda = 0 \implies 2y = 4 \implies y = 2,$$

  and, since $x + y = 1$, we obtain that our *KKT point* is $(-1, 2)$. However, using this point in the first KKT condition, we obtain: $-2 + 4 - 3 + \lambda = 0$, $\lambda = -1$, which is **not a valid value** for λ, so we **discard** this KKT point.

- Case $\alpha, \beta > 0$.
  In this case, we obtain from the KKT conditions that $x = y = 0$, which does not match our initial conditions $x + y = 1$, so no *KKT candidate points* are obtained.

- Case $\alpha > 0, \beta = 0$.
  Looking at our KKT conditions, since $\alpha > 0$, we have that $x = 0$, resulting in $y = 1$ and a possible *KKT point* $(0, 1)$. If we use this point in the second KKT condition, we obtain $0 + 4 + 1 + \lambda = 0$, $\lambda = -5$, which is not a valid value for λ, so we discard again this KKT point.

- Case $\alpha = 0, \beta > 0$.
  Using the same reasoning, we obtain $(1, 0)$ as a possible *KKT point*. We check the λ condition using the first KKT condition: $1 + 0 - 3 + \lambda = 0$, $\lambda = 2$, so we obtain a **KKT factible** point.

Since we have a **unique** candidate, $(1, 0)$, making use of the previously stated theorem, this point is the **minimum**, with a value of $f(1, 0) = -2$.

Problem 3
Let $f : S \subset \mathbb{R}^d \to \mathbb{R}$ be a convex function on the convex set S and we extend it to an $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$ as:

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in S \\ +\infty & \text{if } i \notin S \end{cases}$$

Show that $\tilde{f}$ is a convex function on $\mathbb{R}^d$. Assume that $a + \infty = \infty$ and $a \cdot \infty = \infty$ for $a > 0$.

Let $x, x' \in \mathbb{R}^d$ and consider the segment $\lambda x + (1-\lambda)x'$ with $\lambda in [0,1]$. We can consider two cases:

1. If $x, x' \in S$, since S is convex, $\lambda x + (1-\lambda)x' \in S$ for any $\lambda \in [0,1]$. Also, we know that in S we have $f(x) = \tilde{f}(x)$ and the same happens for x'. Then, for any $\lambda \in [0,1]$:

$$\tilde{f}(\lambda x + (1-\lambda)x') = f(\lambda x + (1-\lambda)x') \overset{(1)}{\leq} \lambda f(x) + (1-\lambda) f(x') = \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(x')$$

   where, in (1) we have used the convexity of f.

2. If $x \in S$ and $x' \notin S$, the convexity inequality is trivial since $\tilde{f}(x') = +\infty$, so, using that we assume that $a \cdot +\infty = +\infty$, we see that for any $\lambda \in [0,1]$

$$\tilde{f}(\lambda x + (1-\lambda)x') \leq \lambda \tilde{f}(x) + (1-\lambda)\tilde{f}(x') = +\infty$$

   The case where $x' \in S$ and $x \notin S$ is analogous.

We have seen that the definition of convexity is fulfilled for any $x, x' \in \mathbb{R}^d$, so $\tilde{f}$ is convex.

Problem 4
Prove **Jensen's inequality** : if f is convex on $\mathbb{R}^d$ and $\sum_1^k \lambda_i = 1$, with $0 \leq \lambda_i \leq 1$, we have for any $x_1, \ldots, x_k \in \mathrm{R}^n$

$$f\left(\sum_1^k \lambda_i x_i\right) \leq \sum_1^k \lambda_i f\left(x_i\right)$$

Hint: just write $\sum_1^k \lambda_i x_i = \lambda_1 x_1 + (1-\lambda_1)\, v$ for an appropriate v and apply repeatedly the definition of a convex function. Start with $k = 3$ and carry on.

Problem 5
Prove that the following function is convex

$$f(x) = \begin{cases} x^2 - 1 & |x| > 1 \\ 0 & |x| \leq 1 \end{cases}$$

and compute its proximal. Which are the fixed points of this proximal?