METHODS OF NONLINEAR OPTIMIZATION: HW#6

(1) Given a C^1 function $g : \mathbb{R}^n \to \mathbb{R}$. How does the level set $L = \{x \in \mathbb{R}^n : g(x) = 0\}$ look like near a point $x^* \in L$ if (i) $\nabla g(x^*) \neq 0$? (ii) $\nabla g(x^*) = 0$? More generally, given a C^1 function $G = [g_1, \ldots, g_p]^T : \mathbb{R}^n \to \mathbb{R}^m$, $(m \leq n)$. How does the constraint set

$$L = \{x \in \mathbb{R}^n : G(x) = 0\} \tag{0.1}$$

look like near a point $x^* \in L$ if (i) $dG(x^*)$ is full rank (i.e. rank m)? (ii) $dG(x^*)$ is rank deficient?

By the local linear approximation of G at x^*,

$$G(x) \approx G(x^*) + dG(x^*)(x - x^*), \quad x \approx x^*.$$

So near x^*, we expect that

$$L = \{x : G(x) = 0\} \approx \{x : \underbrace{G(x^*)}_{=0} + dG(x^*)(x - x^*) = 0\} = x^* + \mathrm{null}(dG(x^*)). \tag{0.2}$$

By the rank-nullity theorem, the latter is exactly an $n - \mathrm{rank}(dG(x^*))$ dimensional plane. Notice also that $\mathrm{null}(dG(x^*))$ is the subspace of $\mathbb{R}^n$ consisting of vectors *orthogonal* to

$$\nabla g_1(x^*), \ldots, \nabla g_m(x^*).$$

Moreover, (0.2) suggests – but not proves – that L is locally a $n - \mathrm{rank}(dG(x^*))$ dimensional surface with the latter as its tangent surface at x^*. ***It turns out that this heuristics is only correct when*** $\mathrm{rank}(dG(x^*)) = m$:

Theorem[Implicit function theorem.][1] If $dG(x^*)$ is full rank, then there exists an injective C^1 function $h : U \overset{\text{open}}{\subset} \mathbb{R}^{n-m} \to \mathbb{R}^n$ such that $L \cap W = h(U)$ for some open ball W in $\mathbb{R}^n$ containing the point x^*. Moreover, the tangent space of L at x^* is given by

$$\mathrm{Image}(dh(x^*)) = \mathrm{null}(dG(x^*)).$$

For a proof of this fact, consult any standard real analysis textbook.

The implicit function theorem/local linear approximation does **not** tell us how L looks like near x^* when $dG(x^*)$ is rank deficient. Consider:

(I) $g(x_1, x_2) = x_1^2 + x_2^2 - 1$, so $x^* := [1/\sqrt{2}, 1/\sqrt{2}]^T \in L$, and $\nabla g(x^*) \neq [0, 0]^T$. How does L look like?

(II) $g(x_1, x_2) = x_1^2 + x_2^2$, so $x^* := [0, 0]^T \in L$, and $\nabla g(x^*) = [0, 0]^T$. How does L look like?

(III) $g(x_1, x_2) = 0$, so $x^* := [0, 0]^T \in L$, and $\nabla g(x^*) = [0, 0]^T$. How does L look like?

Date: June 9, 2020.

[1] The equation "$G(x) = 0$" in (0.1) does not explicitly tell us what points are in L, it tells us how to check if a point is, or is not, in L, namely, $x \in L$ if $G(x) = 0$ and $x \notin L$ if $G(x) \neq 0$. As an example, an explicit representation of a circle is $t \mapsto (\cos(t), \sin(t))$, whereas an implicit representation of the same circle is $x_1^2 + x_2^2 = 1$. The implicit function theorem tells us that, under a regularity condition, the two representations are equivalent locally. How can such an abstract mathematical result be useful in optimization methods? We are about to see that it is instrumental for understanding the Lagrange multiplier theorem.

(5 pts) In each case, does L look like a 0-, 1-, or 2- dimensional surface near x^*?

(10 pts) What do examples (I)-(III) illustrate?

(2) Let's see if we can prove Theorem 11.5 (KKT conditions for **inequality** constrained problems) using the standard Lagrange multiplier theorem in multivariate calculus. For this purpose, we need to first understand how the latter theorem works.

The Lagrange multiplier theorem is for **equality** constrained problems:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad g_i(x) = 0, i = 1, \ldots, m. \tag{0.3}$$

If there are no constraints, optimality implies $\nabla f(x^*) = 0$. Recall that this is equivalent to saying that the directional direction of f at x^* in **any** direction is zero, i.e.

$$f'(x^*; d) = 0, \quad \forall d \in \mathbb{R}^n.$$

When there are constraints, we should not care about directions pointing 'outside' of the constraint surface $L := \{x : h_i(x) = 0, i = 1, \ldots, m\}$; a moment of thought may suggest that the correct optimality condition is

$$f'(x^*; d) = 0, \quad \forall d \text{ tangent to } L \text{ at } x^*. \tag{0.4}$$

Let's be more rigorous. The implicit function theorem offers us an explicit representation $h : U \to \mathbb{R}^n$ of L near x^*, which allows us to express the restriction of f to L near x^* as

$$f \circ h : U \overset{\text{open}}{\subset} \mathbb{R}^{n-m} \to \mathbb{R}.$$

Write $y^* := h^{-1}(x^*)$. As such, x^* is a local minimizer of the constrained optimization problem (0.3) iff y^* is a local minimizer of the *unconstrained* minimization problem $\min_{y \in U} f(h(y))$. But then, by the chain rule, it is necessary that

$$d(f \circ h)(y^*) = df(x^*) \cdot dh(y^*) = 0. \tag{0.5}$$

Assume the linear independence of $\nabla g_1(x^*), \ldots, \nabla g_m(x^*)$.

(10pts) Explain why (0.5) is equivalent to what we expected, namely (0.4).

(10pts) Explain why (0.5) is also equivalent to the Lagrange multiplier condition, namely, there exists $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that

$$\nabla f(x^*) = -\sum_{i=1}^{m} \lambda_i \nabla g_i(x^*).$$

(20 pts) Prove Theorem 11.5 from the Lagrange multiplier condition above. The key issue is how to go from equality to inequality constraints. Hint: think about the effects of the active and inactive constraints.

(5 pts) Do you think the techniques here are enough for establishing Theorem 11.12 on Page 214? Why?

(3) Study Section 12.3.1. It shows that the dual of a linear program (LP) is again a linear program.

(20 pts) Now consider a LP of the form $\min_x c^T x$ s.t. $Ax \geq b$, $x \geq 0$. Derive its dual.[2]

[2] Example 12.13 in the text demonstrates the fact that there is no "one" dual problem for a given primal problem. For LP, however, we do not have this kind of ambiguity.

Recall that the dual of any optimization problem – convex or not – is always a convex optimization problem. This means, in general, the dual of the dual is not the same as the primal.

(20 pts) Prove: for LP, the dual of the dual is always the primal. (You can consider a LP of any form.)

Thomas Yu, Department of Mathematics, Drexel University, 3141 Chestnut Street, 206 Korman Center, Philadelphia, PA 19104, U.S.A.

E-mail address: yut@drexel.edu