Operations Research Techniques and Algorithms (620-361)

Dr Yao-ban Chan

y.chan@ms.unimelb.edu.au
Telephone: 8344 9073
Office: Room 198, Richard Berry Building

Christina Burt
c.burt@ms.unimelb.edu.au
Telephone: 8344 1797
Office: 139 Barry St

Monday 30th April, 2008

Sufficient optimality conditions

As we saw with both unconstrained and equality-constrained problems, stationarity of x^* is generally not enough to ensure that x^* is a local minimum. To be able to deduce that a stationarity point is a local minimum, we need more information.

Specifically, we need to know that

- (NLP) is a *convex program*, or
- a second-order condition holds.

The first of these is dealt with in the following theorem.

Theorem 8: Suppose f and each component function g_j of g is C^1 and convex, h is affine, and a constraint qualification holds for g and h at x^*. Then the KKT conditions hold at x^* if and only if x^* is a local minimum of (NLP).

If (NLP) is not a convex program then we need to use second-order information to establish that x^* is a local minimum. Let (x^*, λ^*, η^*) satisfy the KKT conditions.

Define the *critical cone* at (x^*, λ^*) to be the set

$$\begin{aligned} \mathcal{C}(x^*, \lambda^*) \; := \; \{d \in \Re^n \quad : \quad & \langle \nabla g_i(x^*), d \rangle \leq 0 \text{ if } i \in I(x^*),\ \lambda_i^* = 0, \\ & \langle \nabla g_i(x^*), d \rangle = 0 \text{ if } \lambda_i^* > 0, \\ & \langle \nabla h_j(x^*), d \rangle = 0, \forall j \} \end{aligned}$$

If a constraint is an equality constraint or (at x^*) an active inequality constraint with strictly positive KKT multiplier, the critical cone restricts d to the directions which 'move along' the constraint.

On th other hand, if a constraint is an active inequality constraint with zero KKT multiplier (which acts much like an inactive constraint), then the critical cone restricts d to the directions which 'move towards feasibility'.

Basically, if we move along any direction in the critical cone, the 'feasibleness' of the solution does not change.

Today's Lecture

Second order sufficient conditions

Now observe that the Hessian with respect to x of the Lagrangian function at (x^*, λ^*, η^*) is

$$\nabla^2_{xx} L(x^*, \lambda^*, \eta^*) \ = \ \nabla^2 f(x^*) + \sum_i \lambda_i^* \nabla^2 g_i(x^*) + \sum_j \eta_j^* \nabla^2 h_j(x^*).$$

Remark: If g and h are affine functions, hence $\nabla^2 g(x^*)$ and $\nabla^2 h(x^*)$ are zero, then $\nabla^2_{xx} L(x^*, \lambda^*, \eta^*) = \nabla^2 f(x^*)$.

Let x^* be a KKT point (see Theorem 7). The second-order sufficient condition for x^* to be a local minimum is that $\nabla^2_{xx} L(x^*, \lambda^*, \eta^*)$ is positive definite on the critical cone, that is

$$\text{if} \quad 0 \neq d \in \mathcal{C}(x^*, \lambda^*), \qquad \text{then} \quad d^T \nabla^2_{xx} L(x^*, \lambda^*, \eta^*) d > 0.$$

Of course if $\nabla^2_{xx} L$ is positive definite (over all vectors), then this condition is automatically satisfied!

A slightly weaker second-order sufficient condition is often used. Specifically, x^* is a local minimum if

1. $\lambda_i^* > 0$ for $i \in I(x^*)$, and
2. $\nabla_{xx}^2 L(x^*, \lambda^*, \eta^*)$ is positive definite on the linear subspace

$$\begin{aligned}\{d \in \Re^n \quad : \quad & \langle \nabla g_i(x^*), d \rangle = 0 \text{ if } \lambda_i^* > 0, \\ & \langle \nabla h_j(x^*), d \rangle = 0, \forall j\}.\end{aligned}$$

This condition is much easier to check than the other second-order sufficient condition given above, but is more restrictive since it requires the KKT multipliers λ_i^* associated with active constraints to be strictly positive.

The precise statement of the second order conditions that are sufficient for x^* to be a local minimum is the following:

Theorem 9:
Suppose f, g and h are C^2 functions, and one of the constraint qualifications holds for g and h at x^*. If x^* is stationary and the second-order sufficient condition holds, then x^* is a local minimum of (NLP).

Example. Find a KKT point for the problem $\min_{x \in \Re} x^2 - 2x$ subject to $x \geq 0$. Do the second-order sufficient conditions hold? Repeat this exercise with $x \leq 0$ instead of $x \geq 0$.

Big Example. Consider the NLP

$$\begin{array}{ll} \min & x_1^3 + 4x_2^2 + 16x_3 \\ \text{such that} & x_1 + x_2 + x_3 = 5 \\ & x_1, x_2, x_3 \geq 1. \end{array}$$

Write down the KKT conditions and find all stationary points of the NLP together with their corresponding Lagrange multipliers. At each stationary point, identify the active constraints, write down the critical cone and check a second-order condition. Can you deduce any local minima?

Firstly, we write out the program to give us our constraint functions:

$$\begin{array}{ll} \text{min} & f(x) = x_1^3 + 4x_2^2 + 16x_3 \\ \text{such that} & g_1(x) = 1 - x_1 \leq 0 \\ & g_2(x) = 1 - x_2 \leq 0 \\ & g_3(x) = 1 - x_3 \leq 0 \\ & h(x) = x_1 + x_2 + x_3 - 5 = 0. \end{array}$$

The Lagrangian is

$$\begin{aligned} L(x,\lambda,\eta) &= x_1^3 + 4x_2^2 + 16x_3 \\ &\quad +\lambda_1(1-x_1) + \lambda_2(1-x_2) + \lambda_3(1-x_3) \\ &\quad +\eta(x_1+x_2+x_3-5). \end{aligned}$$

The KKT conditions are

$$\text{KKTa:}\quad \begin{aligned} 3x_1^2 - \lambda_1 + \eta &= 0 \\ 6x_2 - \lambda_2 + \eta &= 0 \\ 16 - \lambda_3 + \eta &= 0 \end{aligned}$$

$$\text{KKTb:}\quad \begin{aligned} &1 - x_1 \le 0, \quad \lambda_1 \ge 0, \quad \lambda_1(1 - x_1) = 0 \\ &1 - x_2 \le 0, \quad \lambda_2 \ge 0, \quad \lambda_2(1 - x_2) = 0 \\ &1 - x_3 \le 0, \quad \lambda_3 \ge 0, \quad \lambda_3(1 - x_3) = 0 \end{aligned}$$

$$\text{KKTc: } x_1 + x_2 + x_3 - 5 = 0$$

To solve this, we take all possible combinations of inequality constraints to be active and see which ones solve the KKT system. Because we have 3 inequality constraints, and any can be active or inactive, this gives us 8 possibilities!

First we test if all 3 constraints can be active. This is easy since it tells us what the values of the decision variables are straight away.

1. If $\lambda_1, \lambda_2, \lambda_3 > 0$ then by KKTb, we have $x_1 = x_2 = x_3 = 1$. But then $x_1 + x_2 + x_3 - 5 = -2 \neq 0$, so KKTc is not satisfied.

Next we test if two constraints can be active. This is also simple because we know two of the x's, and the third follows from the equality constraint.

2. If $\lambda_1 = 0$, $\lambda_2, \lambda_3 > 0$, then by KKTb, $x_2 = x_3 = 1$. From KKTc, $x_1 + 1 + 1 - 5 = 0$ so $x_1 = 3$. Then KKTa gives us $\eta = -3x_1^2 = -27$, and $\lambda_2 = 6x_2 + \eta = -21 \not\geq 0$, so KKTb is not satisfied.
3. If $\lambda_2 = 0$, $\lambda_1, \lambda_3 > 0$, then by KKTb, $x_1 = x_3 = 1$. From KKTc, $x_2 = 3$. Then KKTa gives us $\eta = -6x_2 = -18$, and $\lambda_1 = 3x_1^2 + \eta = -15 \not\geq 0$. So KKTb is not satisfied.

4. If $\lambda_3 = 0$, $\lambda_1, \lambda_2 > 0$, then by KKTb, $x_1 = x_2 = 1$. From KKTc, $x_3 = 3$. Then KKTa gives us $\eta = -16$, and $\lambda_1 = 3x_1^2 + \eta = -13 \not\geq 0$. So KKTb is not satisfied.

The case with one active constraint is a bit harder.

5. If $\lambda_1 = \lambda_2 = 0$, $\lambda_3 > 0$, then by KKTb, $x_3 = 1$. Then from KKTa, $\eta = -3x_1^2 = -6x_2$, which implies that $x_1 = \sqrt{-\frac{\eta}{3}}$ and $x_2 = -\frac{\eta}{6}$. Then from KKTc, $\sqrt{-\frac{\eta}{3}} - \frac{\eta}{6} + 1 - 5 = 0$, which is a quadratic in η and solves to $\eta = -12$. Then $x_1 = 2 \geq 1$ and $x_2 = 2 \geq 1$. Finally $\lambda_3 = 16 + \eta = 4 \geq 0$, so the point $(2, 2, 1)$ is a KKT point.

6. If $\lambda_1 = \lambda_3 = 0$, $\lambda_2 > 0$, then by KKTb, $x_2 = 1$. From KKTa, $\eta = -3x_1^2 = -16$ so $x_1 = \frac{4}{\sqrt{3}}$. KKTc now gives $x_3 = 4 - \frac{4}{\sqrt{3}} \geq 1$. Finally $\lambda_2 = 6x_2 + \eta = -10 \not\geq 0$, so KKTb is not satisfied.
7. If $\lambda_2 = \lambda_3 = 0$, $\lambda_1 > 0$, then by KKTb, $x_1 = 1$. From KKTa, $\eta = -6x_2 = -16$, so $x_2 = \frac{8}{3}$. KKTc now gives us $x_3 = \frac{4}{3}$, and KKTa gives $\lambda_1 = 3x_1^2 + \eta = -13 \not\geq 0$, so KKTb is not satisfied.

It turns out that the case with no active constraints is relatively easy to solve.

8. If $\lambda_1 = \lambda_2 = \lambda_3 = 0$, then KKTa gives us $\eta = -3x_1^2 = -6x_2 = -16$, which implies $x_1 = \frac{4}{\sqrt{3}}$ and $x_2 = \frac{8}{3}$. KKTc now gives $x_3 = \frac{7}{3} - \frac{4}{\sqrt{3}} \not\geq 1$, so KKTb is violated.

Therefore the only KKT point is $(2, 2, 1)$, with multipliers $(0, 0, 4)$ and -12.

Next we check that a constraint qualification holds. This is easy because all constraints are affine. But because we will need it later, we also check the second constraint qualification. At $(2, 2, 1)$ the active constraints are the third inequality constraint and the equality constraint. Therefore the active gradients are

$$\nabla g_3(x) = (0, 0, -1)$$

and

$$\nabla h(x) = (1, 1, 1).$$

It is easy to see that they are linearly independent; furthermore, $d = (0, -1, 1)$ also suffices for Mangasarian-Fromovitz.

To complete our analysis, we show that a sufficient optimality condition holds at $(2, 2, 1)$. Note that this NLP is *not* a convex program because the objective function is not convex. The active gradients give us the critical cone:

$$\mathcal{C}(x^*, \lambda^*) = \{d \in \Re^n : \langle (0, 0, -1), d \rangle = \langle (1, 1, 1), d \rangle = 0\}$$

If $d = (d_1, d_2, d_3)$, then this means that $-d_3 = 0$ and $d_1 + d_2 + d_3 = 0$. This leads to $d_3 = 0$ and $d_2 = -d_1$. So we can express the critical cone as $d_1(1, -1, 0)$ for any $d_1 \in \Re$.

We want to show that the Hessian of the Lagrangian is positive definite on the critical cone. The Hessian is

$$\nabla^2 L(x, \lambda, \eta) = \begin{bmatrix} 6x_1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\nabla^2 L((2,2,1),(0,0,4),-12) = \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which is *not* positive definite (it is positive semi-definite).

But on the critical cone,

$$
\begin{aligned}
& d_1 \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 12 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{bmatrix} d_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\
= \; & d_1^2 \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ -6 \\ 0 \end{bmatrix} \\
= \; & 18 d_1^2 > 0
\end{aligned}
$$

so the Hessian is positive definite on the critical cone.

In summary:

$(2, 2, 1)$ is a local minimum of the non-linear program, with KKT multipliers $(0, 0, 4)$ and -12.