# 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 14th April, 2008

# Today's Lecture

Constrained optimisation
- Optimality conditions

We now check the constraint qualifications at each of these points. Since in this example, $h$ is *not* affine, ($h_1$ in particular is not affine), the first does not hold, so we look at the constraint gradients at each point. We have

$$\nabla h_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \\ 0 \end{bmatrix} \quad \text{and} \quad \nabla h_2(x) = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}.$$

The second constraint qualification is satisfied if these vectors are linearly independent, that is if the matrix

$$\nabla h(x) = \begin{bmatrix} 2x_1 & 2 \\ 2x_2 & 2 \\ 0 & 1 \end{bmatrix}$$

has full column rank, that is has rank 2.

Now

$$\nabla h\left(\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 1+2\sqrt{2} \end{bmatrix}\right) = \begin{bmatrix} -\sqrt{2} & 2 \\ -\sqrt{2} & 2 \\ 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -\sqrt{2} \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

which clearly has rank 2. So the constraint qualification holds at $x = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 1+2\sqrt{2})^T$.

Similarly

$$\nabla h\left(\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 1-2\sqrt{2} \end{bmatrix}\right) = \begin{bmatrix} \sqrt{2} & 2 \\ \sqrt{2} & 2 \\ 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \sqrt{2} \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

which also clearly has rank 2. So the constraint qualification holds at $x = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 1-2\sqrt{2})^T$ also.

Thus, by Theorem 4, (and since $f$ and $h$ are obviously $C^1$), if the NLP has a locally optimal point, it must be one of these two points. To find out more, we need to look at second-order information.

**Example.**

Write down and solve the first-order necessary conditions for the equality-constrained NLP when $n = 2$, $f(x) = x_1^2 + x_2^2/4$, and $h(x) = x_1 + x_2 + 1$.

We prove Theorem 4 for the case $n = 2$, $q = 1$ (one constraint) more formally.

The feasible region must be able to be expressed as a parametric curve $x(t)$, where $x \in \Re^n$ and $t$ is a single variable. However, the feasible region is defined as the region $h(x) = 0$. Therefore

$$h(x(t)) = 0 \text{ for all } t.$$

Differentiating this expression by $t$ using the multi-dimensional chain rule gives us

$$\begin{aligned}\frac{d}{dt}h(x(t)) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i} h(x(t)) \frac{d}{dt} x_i(t) \\ &= \langle \nabla h(x(t)), x'(t) \rangle \\ &= 0.\end{aligned}$$

This shows us that $\nabla h(x(t))$ is orthogonal to $x'(t)$. In particular, if $x^*$ is a local minimum and corresponds to $t^*$ on the parametric curve, we know that $\nabla h(x^*)$ is orthogonal to $x'(t^*)$.

However, because $x^*$ is a local minimum, the one-dimensional function

$$q(t) = f(x(t)),$$

which is the function 'sliced' along the feasible region, achieves a minimum at $t^*$. This implies that $q'(t^*) = 0$. However, again from the chain rule,

$$\begin{aligned} q'(t) &= \frac{d}{dt} f(x(t)) \\ &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(x(t)) \frac{d}{dt} x_i(t) \\ &= \langle \nabla f(x(t)), x'(t) \rangle. \end{aligned}$$

This implies that $\langle \nabla f(x(t^*)), x'(t^*) \rangle = 0$, or that $\nabla f(x(t^*))$ and $x'(t^*)$ are orthogonal. But $x'(t^*)$ is also orthogonal to $\nabla h(x^*)$! Therefore $\nabla h(x^*)$ and $\nabla f(x(t^*))$ are parallel (remember we are in two dimensions). This gives

$$\nabla f(x(t^*)) \propto \nabla h(x^*)$$

$$\nabla f(x(t^*)) = -\eta^* \nabla h(x^*)$$

$$\nabla f(x(t^*)) + \eta^* \nabla h(x^*) = 0$$

which, again, is Lagrange's condition.

As in unconstrained optimization, a stationary point need not be a local or global minimum. However, if we know more, we can conclude that a stationary point is a local or global minimum.

For example, if the objective and constraint functions have particular convexity properties, or if particular second-order properties hold at the stationary point, we can conclude that the stationary point is a global or local minimum respectively. The first of these results is stated in the following theorem.

**Theorem 5:** *If $f$ is $C^1$ and convex, and $h$ is affine, then a point $x^*$ is stationary for the equality constrained NLP if and only if it is a global minimum.*

This is comparable to the unconstrained result which requires $f$ to be $C^1$ and convex.
If the objective function is nonconvex and/or the constraints are nonlinear, then the above result does not hold. In this case we rely on a second order sufficient condition for a stationary point to be a local minimum. This is given in the following theorem.

Let $x^*$ be a stationary point, with a corresponding Lagrange multiplier $\eta^*$. Let $\nabla^2_{xx} L(x^*, \eta^*)$ be the Hessian of the Lagrangian with respect to $x$, that is

$$\nabla^2_{xx} L(x^*, \eta^*) \ = \ \nabla^2 f(x^*) + \sum_{j=1}^{q} \eta_j^* \nabla^2 h_j(x^*).$$

Define the set of feasible directions at $x^*$ to be

$$\mathcal{C}(x^*) = \{d \in \Re^n : d \neq 0, \nabla h(x^*)^T d = 0\}.$$

$\mathcal{C}(x^*)$ is the nullspace (or kernel) of the transpose of the Jacobian $\nabla h(x^*)$.

The next theorem provides a second-order sufficient condition for $x^*$ to be a local minimum.

**Theorem 6:** *If $\nabla^2_{xx} L(x^*, \eta^*)$ is positive definite on $\mathcal{C}(x^*)$, that is*

$$\textit{if} \quad 0 \neq d \in \Re^n,\ \nabla h(x^*)^T d = 0, \quad \textit{then} \quad d^T \nabla^2_{xx} L(x^*, \eta^*) d \, > \, 0,$$

*then $x^*$ is a local minimum.*

This is analogous to the unconstrained second-order condition which requires $\nabla^2 f(x^*)$ to be positive definite.

Theorem 6 tells us that if the Hessian of the Lagrangian function with respect to the $x$ variables is positive definite in directions which maintain feasibility (in a local sense, at least), then the stationary point must be a local optimum (assuming that $f$ and $h$ are $C^2$).

Let's continue with our previous example:

$$\begin{array}{llcccccccc}\min & f(x) = & 2x_1^2 + 2x_2^2 + 4x_1x_2 + x_1x_3 + x_2x_3 \\ \text{s.t.} & h_1(x) = & x_1^2 & + & x_2^2 & & & -1 & = 0 \\ & h_2(x) = & 2x_1 & + & 2x_2 & + & x_3 & -1 & = 0.\end{array}$$

The first sufficient condition (requiring $f$ to be convex and $h$ to be affine) does not apply, as $h$ is not affine.

Let us try the second condition. $f$ and $h$ are obviously $C^2$. Also we have

$$\nabla^2_{xx} L(x, \eta) = \begin{bmatrix} 4 + 2\eta_1 & 4 & 1 \\ 4 & 4 + 2\eta_1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Let us consider the stationary point $x^* = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 1 + 2\sqrt{2} \end{bmatrix}$ with

Lagrange multipliers $\eta^* = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \sqrt{2} \end{bmatrix}$.

The Hessian of the Lagrangian at this point is

$$\nabla^2_{xx} L(x^*, \eta^*) = \begin{bmatrix} 4+\sqrt{2} & 4 & 1 \\ 4 & 4+\sqrt{2} & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

We wish to know if this is positive definite with respect to the directions which maintain feasibility. This is the nullspace of the transpose of the Jacobian

$$\nabla h(x) = \begin{bmatrix} 2x_1 & 2 \\ 2x_2 & 2 \\ 0 & 1 \end{bmatrix}.$$

Now

$$\nabla h(x^*)^T d = 0 \quad \Rightarrow \quad \begin{bmatrix} -\sqrt{2} & -\sqrt{2} & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad \left.\begin{array}{c} -\sqrt{2}d_1 - \sqrt{2}d_2 = 0 \\ 2d_1 + 2d_2 + d_3 = 0 \end{array}\right\} \quad \Rightarrow \quad d_2 = -d_1 \quad \text{and} \quad d_3 = 0$$

$$\Rightarrow \quad d = \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix}$$

for some $d_1 \in \Re$.

Thus the directions of interest are given by

$$\mathcal{C}(x^*) = \{d \in \Re^n \ : \ d = \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix}, \text{ for some } d_1 \neq 0\}.$$

We need to check if $d^T \nabla^2_{xx} L(x^*, \eta^*) d > 0$ for all $d \in \mathcal{C}(x^*)$; if so, we can then apply Theorem 6 to deduce that $x^*$ is a local optimum.

For $d_1 \in \Re$, $d_1 \neq 0$, we have

$$\begin{aligned}
& (d_1, -d_1, 0)\nabla^2_{xx} L(x^*, \eta^*) \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix} \\
= \ & (d_1, -d_1, 0) \begin{bmatrix} 4+\sqrt{2} & 4 & 1 \\ 4 & 4+\sqrt{2} & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix} \\
= \ & (d_1, -d_1, 0) \begin{bmatrix} \sqrt{2}d_1 \\ -\sqrt{2}d_1 \\ 0 \end{bmatrix} \\
= \ & 2\sqrt{2}d_1^2 \\
> \ & 0.
\end{aligned}$$

Thus $\nabla^2_{xx} L(x^*, \eta^*)$ is positive definite on $\mathcal{C}(x^*)$, so the second order sufficiency condition does holds. Therefore the point $x^* = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 1 + 2\sqrt{2})$ is a local minimum, with Lagrange multipliers $\eta^* = (\frac{1}{\sqrt{2}}, \sqrt{2})$.

What about $x^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 1 - 2\sqrt{2})$?

The Hessian of the Lagrangian at this point is

$$\nabla^2_{xx} L(x^*, \eta^*) = \begin{bmatrix} 4-\sqrt{2} & 4 & 1 \\ 4 & 4-\sqrt{2} & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

The transpose of the Jacobian at this point is

$$\nabla h(x^*)^T = \begin{bmatrix} \sqrt{2} & \sqrt{2} & 0 \\ 2 & 2 & 1 \end{bmatrix}.$$

Now

$$\nabla h(x^*)^T d = 0 \quad \Rightarrow \quad \begin{bmatrix} \sqrt{2} & \sqrt{2} & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad \left. \begin{array}{c} \sqrt{2}d_1 + \sqrt{2}d_2 = 0 \\ 2d_1 + 2d_2 + d_3 = 0 \end{array} \right\} \quad \Rightarrow \quad d_2 = -d_1 \quad \text{and} \quad d_3 = 0$$

$$\Rightarrow \quad d = \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix}$$

for some $d_1 \in \Re$.

Thus the directions of interest are given by

$$\mathcal{C}(x^*) = \{d \in \Re^n \ : \ d = \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix}, \text{ for some } d_1 \neq 0\}.$$

We need to check if $d^T \nabla^2_{xx} L(x^*, \eta^*) d > 0$ for all $d \in \mathcal{C}(x^*)$; if so, we can then apply Theorem 6 to deduce that $x^*$ is a local optimum.

For $d_1 \in \Re$, $d_1 \neq 0$, we have

$$\begin{aligned}
& (d_1, -d_1, 0)\nabla^2_{xx} L(x^*, \eta^*) \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix} \\
= \;& (d_1, -d_1, 0) \begin{bmatrix} 4 - \sqrt{2} & 4 & 1 \\ 4 & 4 - \sqrt{2} & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ -d_1 \\ 0 \end{bmatrix} \\
= \;& (d_1, -d_1, 0) \begin{bmatrix} -\sqrt{2} d_1 \\ \sqrt{2} d_1 \\ 0 \end{bmatrix} \\
= \;& -2\sqrt{2} d_1^2 \\
< \;& 0.
\end{aligned}$$

Thus $\nabla^2_{xx} L(x^*, \eta^*)$ is certainly *not* positive definite on $\mathcal{C}(x^*)$, so the second order sufficiency condition does *not* hold, and we cannot deduce that this point is a local optimum. (In fact, $\nabla^2_{xx} L(x^*, \eta^*)$ is *negative* definite on $\mathcal{C}(x^*)$ so $x^*$ is a local maximum.)

Putting this together with the fact that the only other stationary point is a minimum, we deduce that the global minimum of the function lies at the point $(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 1 + 2\sqrt{2})$.

**Example.** Check the second-order conditions for the earlier problem: $n = 2$, $f(x) = x_1^2 + x_2^2/4$, and $h(x) = x_1 + x_2 + 1$.