Optimality conditions for nonlinear constraints (12/6)

Constrained problem with nonlinear equality constraints:

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g(x) = 0 \end{aligned} \qquad (*)$$

where $f: \mathbb{R}^n \to \mathbb{R}$, and $g: \mathbb{R}^n \to \mathbb{R}^m$, $m < n$.

And with inequality constraints:

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g(x) \geq 0. \end{aligned} \qquad (**)$$

Def: a point x is regular for $(*)$ if $Dg(x)$ is full rank.
Equivalently, the set $\{\nabla g_i(x) : 1 \leq i \leq m\}$ is linearly independent.

Example: Two equality constraints:

$$g_1(x,y,z) = x^2 + y^2 + z^2 - 3 = 0 \quad \leftarrow \text{a sphere}$$

$$g_2(x,y,z) = 2x - 4y + z^2 + 1 = 0$$

The gradients:

$$\nabla g_1(x,y,z) = (2x, 2y, 2z)$$

$$\nabla g_2(x,y,z) = (2, -4, 2z)$$

The point $(x,y,z) = (1,1,1)$ is feasible. At this point:

$$\nabla g_1(1,1,1) = (2,2,2)$$

$$\nabla g_2(1,1,1) = (2,-4,2)$$

Hence $\{\nabla g_1(1,1,1), \nabla g_2(1,1,1)\}$ are linearly dependent and $(1,1,1)$ is regular.

<u>Def</u>: A point x is regular for $(**)$ if the set

$$\{\nabla g_i(x) \mid g_i(x) = 0\} \quad (\leftarrow \text{the gradients of the active constraints})$$

is linearly independent.

<u>Example</u>: Consider the single inequality constraint:

$$g(x,y) = \left[\frac{x^2}{2} + \frac{y^2}{2} - 1\right]^3 \geq 0.$$

Note that g is <u>radial</u>: i.e., if $r^2 = x^2 + y^2$, then:

$$g(x,y) = g(r) = \left(\frac{r^2}{2} - 1\right)^3 \geq 0.$$

We have:

$$\frac{\partial g}{\partial x} = 3x\left(\frac{x^2}{2} + \frac{y^2}{2} - 1\right)^2 = 3x\left(\frac{r^2}{2} - 1\right)^2$$

$$\frac{\partial g}{\partial y} = 3y\left(\frac{x^2}{2} + \frac{y^2}{2} - 1\right)^2 = 3y\left(\frac{r^2}{2} - 1\right)^2$$

So, if $r = \sqrt{2}$, then $\nabla g = 0$. But the set of points for which $r = \sqrt{2}$, is exactly the set $\{(x,y) : g(x,y) = 0\}$, i.e. which are on the boundary of the region defined by the inequality. So, there are no irregular points for this inequality constraint.

Note that this example shows us something important. We only care about the set defined by a constraint. There are always many different functions which realize the same region (e.g. multiplying by a positive constant leaves the region unchanged).

For this example, we could use the much simpler function

$$g(x,y) = \frac{r^2}{2} - 1 = \frac{x^2}{2} + \frac{y^2}{2} - 1$$

to realize the same set. But for this choice, all points on the boundary are regular:

$$\frac{\partial g}{\partial x} = x, \quad \frac{\partial g}{\partial y} = y \quad \Rightarrow \nabla g \neq 0 \text{ if } r = \sqrt{2}.$$

Let's look at the optimality conditions for $(*)$. Recall that the Lagrange function for $(*)$ is:

$$L(x, \lambda) = f(x) - \lambda^T g(x),$$

where $\lambda \in \mathbb{R}^m$.

Let's also define $Z(x)$ to be a basis matrix for the nullspace of $\nabla g(x)$.

Theorem: Nec. conds for (*)

Let x^* be constrained local min for (*). If x^* is regular, then there exists λ^* such that:

- $$\left.\frac{\partial L}{\partial x}\right|_{(x^*,\lambda^*)} = 0 \iff Z(x^*)^T \nabla f(x^*) = 0$$
- $$Z(x^*)^T \left.\frac{\partial^2 L}{\partial x \partial x^T}\right|_{(x^*,\lambda^*)} Z(x^*)$$ is pos. semidefinite

Theorem: Suf. conds for (*)

Let x^* be feasible. If λ^* exist st:

- $$\left.\frac{\partial L}{\partial x}\right|_{(x^*,\lambda^*)} = 0$$
- $$Z(x^*)^T \left.\frac{\partial^2 L}{\partial x \partial x^T}\right|_{(x^*,\lambda^*)} Z(x^*)$$ is pos definite

then x^* is a strict constrained local minimizer.

Compare w/ earlier results. If g is linear, then $Dg \equiv A$.

Hence, $Z(x^*) \equiv Z$, and:

$$0 = \left.\frac{\partial L}{\partial x}\right|_{(x^*,\lambda^*)} = \nabla f(x^*) - A^T\lambda^* \Rightarrow \nabla f(x^*) = A^T\lambda^*.$$

Example: Solve:

$$\text{minimize} \quad x^2 - y^2$$
$$\text{subject to} \quad x^2 + 2y^2 - 4 = 0$$

Lagrange function is:

$$L(x, y, \lambda) = x^2 - y^2 - \lambda(x^2 + 2y^2 - 4)$$

And its gradient is:

$$\frac{\partial L}{\partial x} = 2x - 2\lambda x, \qquad \frac{\partial L}{\partial y} = -2y - 4\lambda y$$

Hence, need:

$$2x^* - 2\lambda^* x^* = 0$$
$$-2y^* - 4\lambda^* y^* = 0$$

For first equation, two options: $x^* = 0$ or $\lambda^* = 1$.

- If $x^* = 0$, then $\lambda^* = -\frac{1}{2}$, and from feasibility, $y^* = \pm\sqrt{2}$.
- If $\lambda^* = 1$, then $y^* = 0$ and $x^* = \pm 2$ from feasibility.

So, four stationary points:

x^*	y^*	λ^*
0	$\sqrt{2}$	$-\frac{1}{2}$
0	$-\sqrt{2}$	$-\frac{1}{2}$
2	0	1
-2	0	1

To find the minimizers, need to look at the reduced Hessian:

$$\left.\begin{array}{l} \frac{\partial^2 L}{\partial x^2} = 2 - 2\lambda \\ \frac{\partial^2 L}{\partial x \partial y} = 0 \\ \frac{\partial^2 L}{\partial y^2} = -2 - 4\lambda \end{array}\right\} \Rightarrow \nabla^2 L = \begin{bmatrix} 2 - 2\lambda & 0 \\ 0 & -2-4\lambda \end{bmatrix}$$

Have to check each stationary points individually. For the reduced Hessian, need the basis matrix $Z(x^*, y^*)$ for the nullspace of $Dg(x^*, y^*)$. Note that $\nabla g(x,y) = (2x, 4y)$.

$(x^*, y^*) = (0, \sqrt{2})$:

$$Dg(0, \sqrt{2}) = [0, 4\sqrt{2}] \Rightarrow Z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Since $\lambda^* = -\frac{1}{2}$, have $\nabla^2 L = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$, so:

$$Z^T \nabla^2 L Z = [1 \ \ 0] \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 3 > 0$$

Hence, this is a strict local minimum.

$(x^*, y^*) = (0, -\sqrt{2})$:

Also a strict local minimum by almost same reasoning.

Can also check remaining two cases using similar approach.

Theorem: Nec. conds for $(**)$

Let x^* be a local minimum for $(**)$.
Let $Z(x^*)$ be a basis matrix for the nullspace of the Jacobian of the active constraints. If x^* is regular, then there exists λ^* such that:

- $\left.\frac{\partial L}{\partial x}\right|_{(x^*,\lambda^*)} = 0 \iff Z(x^*)^T \nabla f(x^*) = 0$
- $\lambda^* \geq 0$
- $\lambda^{*T} g(x^*) = 0$
- $Z(x^*)^T \left.\frac{\partial^2 L}{\partial x \partial x^T}\right|_{(x^*,\lambda^*)} Z(x^*)$ is pos. semidef

These are the KKT conditions, which we saw before.

Theorem: Suff. conds for $(**)$:

Let x^* be feasible and suppose λ^* exists st:

- $\left.\frac{\partial L}{\partial x}\right|_{(x^*,\lambda^*)} = 0$
- $\lambda^* \geq 0$
- $\lambda^{*T} g(x^*) = 0$
- $Z_+(x^*)^T \frac{\partial^2 L}{\partial x \partial x^T} Z_+(x^*)$ is pos def.

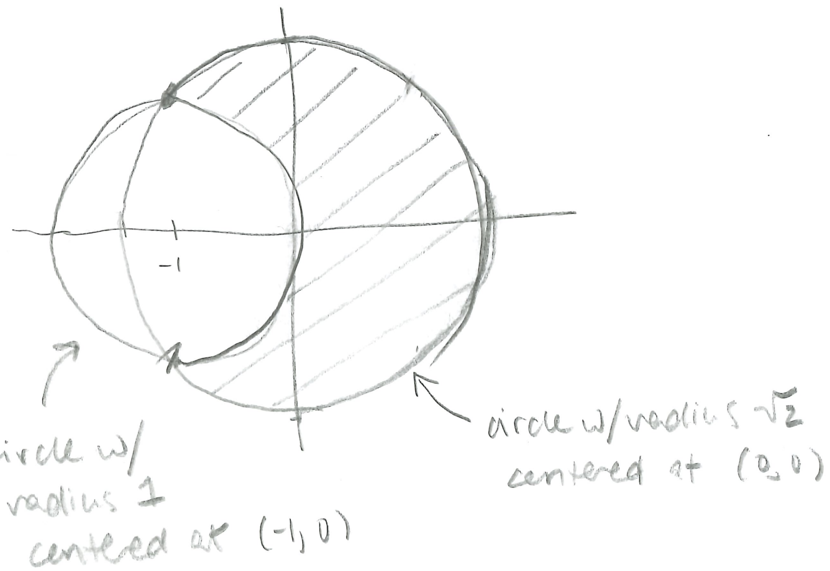
Then, x^* is a strict constrained local minimizer for $(**)$.

where $Z_+(x^*)$ is a basis matrix for the nondegenerate constraints at x^* (active & nonzero Lagrange multipliers).

Example: Solve:

$$\begin{aligned} &\text{minimize} && x \\ &\text{subject to} && (x+1)^2 + y^2 \geq 1 \\ & && x^2 + y^2 \leq 2 \end{aligned}$$

Constraints define this region:



By symmetry, it's clear that $(-1, \pm 1)$ are the two minimizers. Procedure is use KKT theorems above to check this. Procedure is similar to previous example.