LECTURE 18

MATH 409/509

SPRING 2011

NONLINEAR OPTIMIZATION WITH EQUALITY CONSTRAINTS (PART IV)

METHOD OF MULTIPLIERS

A simple approach to solve (NEP) based on application of Newton's method to find a root of $\nabla\mathcal{L}(x,\lambda)$.

$\boxed{\text{Newton Iteration on } \nabla\mathcal{L}(x,y)}$

(1) Solve

$$\underbrace{\nabla^2\mathcal{L}(x_k,\lambda_k)}_{(n+m)\times(n+m)}\begin{bmatrix}\Delta x_k\\ \Delta\lambda_k\end{bmatrix} = -\underbrace{\nabla\mathcal{L}(x_k,\lambda_k)}_{\text{in } \mathbb{R}^{n+m}}$$

for $(\Delta x_k, \Delta\lambda_k)$.

(2) $$\begin{bmatrix}x_{k+1}\\ \lambda_{k+1}\end{bmatrix} = \begin{bmatrix}x_k\\ \lambda_k\end{bmatrix} + \alpha_k\begin{bmatrix}\Delta x_k\\ \Delta\lambda_k\end{bmatrix}$$

where α_k is determined by an inexact line search.

Note that in (1)

$$\nabla^2 \mathcal{L}(x,\lambda) = \begin{bmatrix} \overbrace{\nabla^2 f(x) - \sum_{j=1}^{m} \lambda_j \nabla^2 c_j(x)}^{\nabla^2_{xx}\mathcal{L}} & \overbrace{-[\nabla c_1(x) \dots \nabla c_m(x)]}^{\nabla^2_{x\lambda}\mathcal{L}} \\ \underbrace{-J(x)}_{\nabla^2_{\lambda x}\mathcal{L}} & \underbrace{0}_{\nabla^2_{\lambda\lambda}\mathcal{L}} \end{bmatrix}$$

$$= \begin{bmatrix} \nabla^2 f(x) - \sum_{j=1}^{m} \lambda_j \nabla^2 c_j(x) & -J(x)^T \\ -J(x) & 0 \end{bmatrix}$$

Determine α_k by means of a backtracking line-search with the sufficient decrease condition

$$\|\nabla \mathcal{L}(x_{k+1}, \lambda_{k+1})\| \leqslant (1-\alpha_k \mu)\, \|\nabla \mathcal{L}(x_k, \lambda_k)\|$$

where $\mu \in (0,1)$ is fixed.

EXAMPLE

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad x^T x$$

subject to

$$Ax = b$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$

$$\mathcal{L}(x,\lambda) = x^T x - \lambda^T(Ax-b)$$

$$\nabla \mathcal{L}(x,\lambda) = \begin{bmatrix} 2x - A^T\lambda \\ -Ax+b \end{bmatrix}$$

$$\nabla^2 \mathcal{L}(x,\lambda) = \begin{bmatrix} 2I & -A^T \\ -A & 0 \end{bmatrix}$$

Choose $(\Delta x_k, \Delta \lambda_k)$ at (x_k, λ_k) such that

$$\begin{bmatrix} 2I & -A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix} = \begin{bmatrix} 2x_k - A^T\lambda_k \\ -Ax_k+b \end{bmatrix}$$

<u>ALGORITHM</u> (Method of Multipliers)

Given $x_0 \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}^m$, $\mu \in (0,1)$ and $k=0$

While $\|\nabla \mathcal{L}(x_k,\lambda_k)\| > \epsilon$

(1) Solve the system

$$\nabla^2 \mathcal{L}(x_k,\lambda_k) \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix} = -\nabla \mathcal{L}(x_k,\lambda_k)$$

(2) Perform backtracking line-search

$$\alpha_k = 1$$

While $\left(\|\nabla \mathcal{L}(x_k + \alpha_k \Delta x_k, \lambda_k + \alpha_k \Delta \lambda_k)\| > (1 - \alpha_k \mu) \|\nabla \mathcal{L}(x_k, \lambda_k)\| \right)$

$$\alpha_k = \alpha_k / 2$$

end

(3) $$\begin{bmatrix} x_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} + \alpha_k \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix}$$

$$k = k+1$$

end

Guaranteed to converge only to a stationary point of $\mathcal{L}(x, \lambda)$.

CONSTRAINT QUALIFICATIONS

Recall that

$$T^\circ \text{ at } x_* \subseteq \text{Null}(J(x_*))$$

A constraint qualification ensures

$$(*) \quad T^\circ \text{ at } x_* \supseteq \text{Null}(J(x_*))$$

-LCCQ-

THM (Linear-Constraints Constraint Qualification)

Suppose that the constraints $c_i(x)$ are linear, that is

$$c(x) = Ax - b$$

for some $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Then the constraint qualification holds at all $x_* \in \mathbb{F}$.

PROOF

We only need to show inclusion $(*)$.

Since $J(x_*) = A$, for all $\underset{p\neq 0}{p} \in Null(J(x_*))$ we have

$$Ap = 0.$$

Consider the path

$$x(\alpha) = x_* + \alpha p,$$

which is feasible, since

(i) $x(0) = x_*$

(ii) $x'(0) = p \neq 0$

(iii) $c(x(\alpha)) = A(x_* + \alpha p) - b$

$$= \underbrace{(Ax_* - b)}_{0,\ x_* \in \mathbb{F}} + \alpha\,(\underbrace{Ap}_{0})$$

$$= 0 \quad \text{for all } \alpha$$

Consequently $x'(0) = p \in (T^\circ \text{ at } x_*)$
implying

$$\text{Null}(J(x_*)) \subseteq T^\circ \text{ at } x_*$$

$\square$

EXAMPLE

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad x^T x$$
$$\text{subject to}$$
$$Ax = b$$

By LCCQ the constraint qualification holds at all $x_* \in \mathbb{F}$. Consequently

x_* is a local minimizer

$\Longrightarrow$

(i) $Ax_* = b$ and (ii) $2x_* = A^T \lambda$
for some $\lambda \in \mathbb{R}^n$

THM (Linear Independence Constraint Qualification) – LICQ –
Suppose $x_* \in \mathbb{F}$ is such that the set
$$\{\nabla c_1(x_*), \dots, \nabla c_m(x_*)\}$$
is linearly independent. Then the constraint qualification holds at x_*.

· PROOF

Let $p \in \mathrm{Null}(J(x_*))$ and Z be a matrix whose columns form a basis for $\mathrm{Null}(J(x_*))$. Define also p_z such that $p = Z p_z$.

Consider a path $x(\alpha)$ as a solution of the differential equation

$$x'(\alpha) = Z(x)\, p_z, \qquad x(0) = x_*$$

where $Z(x)$ is a matrix whose columns form a ~~orthonorm~~ basis for $\mathrm{Null}(J(x))$.

Since $Z(x)$ is a continuous function of x (due to the assumption $J(x)$ has linearly independent rows, equivalently $J(x)$ has full rank), the differential equation above has a solution.

· The path $x(\alpha)$ is feasible, since

$$c_j(x(\alpha)) = c_j(x(0)) + \int_0^\alpha c_j'(x(t))\,dt$$

$$= c_j(x_*) + \int_0^\alpha \nabla c_j(x(t))^T x'(t)\,dt$$

$$= \underbrace{c_j(x_*)}_{0} + \int_0^\alpha \underbrace{\nabla c_j(x(t))^T Z(x(t)) p_Z}_{0}\,dt$$

$$= 0$$

Note that above we use the fact that

$$J(x(t))\, Z(x(t)) = 0$$

$$\Longrightarrow$$

$$\nabla c_j(x(t))^T Z(x(t)) = 0$$

We deduce $x'(0) = Z p_Z = p \in (T^\circ \text{ at } x_*)$
implying

$$\text{Null}(J(x_*)) \subseteq T^\circ \text{ at } x_* \qquad \square$$

EXAMPLE

$$\underset{x \in \mathbb{R}^3}{\text{minimize}} \quad x_1 + x_2 + x_3$$

subject to

$$x_1^2 + x_2^2 + x_3^2 = 1$$

$$x_1^2 + 3x_2^2 = 1$$

Constraint gradients

$$\{\nabla c_1(x), \nabla c_2(x)\} = \left\{ \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{bmatrix}, \begin{bmatrix} 2x_1 \\ 6x_2 \\ 0 \end{bmatrix} \right\}$$

are linearly independent if $x_3 \neq 0$, $x_2 = 0$.

Suppose $x_* \in \mathbb{F}$ such that $(x_*)_3 \neq 0$ and $(x_*)_2 \neq 0$. Then

$$x_* \text{ is a local minimizer}$$

$$\Longrightarrow$$

$$\text{(i)} \quad (x_*)_1^2 + (x_*)_2^2 + (x_*)_3^2 = 1$$

$$(x_*)_1^2 + 3(x_*)_2^2 = 1$$

and

$$\text{(ii)} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2(x_*)_1 & 2(x_*)_1 \\ 2(x_*)_2 & 6(x_*)_2 \\ 2(x_*)_3 & 0 \end{bmatrix} \lambda$$

for some $\lambda \in \mathbb{R}^2$