

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 $\nabla \mathcal{L}(x, \lambda)$.

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)

$$\nabla^2 \mathcal{L}(x, \lambda) = \begin{bmatrix} \underbrace{\nabla_{xx}^2 \mathcal{L}}_{\nabla_f^2(x) - \sum_{j=1}^m \lambda_j \nabla_{Cj}^2(x)} & \underbrace{-[\nabla_{C_1}(x) \dots \nabla_{C_m}(x)]^T}_{0} \\ \underbrace{-J(x)}_{\nabla_{\lambda x}^2 \mathcal{L}} & \underbrace{\nabla_{\lambda \lambda}^2 \mathcal{L}}_{0} \end{bmatrix}$$

$$= \begin{bmatrix} \nabla_f^2(x) - \sum_{j=1}^m \lambda_j \nabla_{Cj}^2(x) & -J(x)^T \\ -J(x) & 0 \end{bmatrix}$$

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

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

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

EXAMPLE

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

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

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

ALGORITHM (Method of Multipliers)

Given $x_0 \in \mathbb{R}^n$, $\lambda_0 \in \mathbb{R}^m$, $M \in (0, 1)$ 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 $\| \nabla \mathcal{L}(x_k + \alpha_k \Delta x_k, \lambda_k + \alpha_k \Delta \lambda_k) \| > (1 - \alpha_k M) \| \nabla \mathcal{L}(x_k, \lambda_k) \|$

$$\alpha_k = \alpha_k / 2$$

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

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

A constraint qualification ensures

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

- 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 $x_* \in F$.

PROOF

We only need to show inclusion (*). Since $J(x_*) = A$, for all $p \neq 0 \in \text{Null}(J(x_*)$ we have

$$Ap = 0.$$

Consider the path

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

which is feasible, since

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

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

$$\begin{aligned} (iii) \quad c(x(\alpha)) &= A(x_* + \alpha p) - b \\ &= (\underbrace{Ax_* - b}_0) + \alpha (\underbrace{Ap}_0) \\ &= 0 \quad \text{for all } \alpha \end{aligned}$$

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

implying

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

□

EXAMPLE

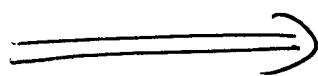
$$\begin{array}{ll} \text{minimize} & x^T x \\ x \in \mathbb{R}^n & \end{array}$$

subject to

$$Ax = b$$

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

x_* is a local minimizer



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

THM (Linear Independence Constraint Qualification)
Suppose $x_* \in F$ 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 $p \in \text{Null}(\mathcal{J}(x_*))$ and Z be a matrix whose columns form a basis for $\text{Null}(\mathcal{J}(x_*))$. Define also p_Z such that $p = Z p_Z$.

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

$$x'(\alpha) = Z(x) p_Z, \quad x(0) = x_*$$

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

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

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

$$\begin{aligned}
 c_j(x(\alpha)) &= c_j(x(0)) + \int_0^\alpha c_j'(x(t)) dt \\
 &= c_j(x_*) + \int_0^\alpha \nabla c_j(x(t))^T x'(t) dt \\
 &= \underbrace{c_j(x_*)}_0 + \int_0^\alpha \underbrace{\nabla c_j(x(t))^T Z(x(t)) p_z}_{0} dt \\
 &= 0
 \end{aligned}$$

Note that above we use the fact that

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

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

We deduce $x'(0) = Z p_z = p \in T^*_{x_*}$

implying

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

□

EXAMPLE

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

subject to

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

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

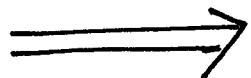
Constraint gradients

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

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

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

x_* is a local minimizer



$$(i) \quad (x_*)_1^2 + (x_*)_2^2 + (x_*)_3^2 = 1$$
$$(x_*)_1^2 + 3(x_*)_2^2 = 1$$

and

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

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