10. NonLinear Optimization

春笋

NonLinear Optimization:both the objective function and the constraints are nonlinear

- penalty- and barrier-function methods(内点法)
- gradient projection methods
- sequential quadratic-programming（SQP）methods

SQP with equality constraints

$$\text{minimize } f(\mathbf{x}) \tag{15.1a}$$

$$\text{subject to:} \quad a_i(\mathbf{x}) = 0 \quad \text{for } i = 1, 2, \ldots, p \tag{15.1b}$$

Lagrangian

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{p} \lambda_i a_i(\mathbf{x})$$

$$\nabla \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$$

$$\{\mathbf{x}_{k+1}, \boldsymbol{\lambda}_{k+1}\} = \{\mathbf{x}_k + \boldsymbol{\delta}_x, \boldsymbol{\lambda}_k + \boldsymbol{\delta}_\lambda\}$$

$$\nabla \mathcal{L}(\mathbf{x}_{k+1}, \boldsymbol{\lambda}_{k+1}) \approx \nabla \mathcal{L}(\mathbf{x}_k, \boldsymbol{\lambda}_k) + \nabla^2 \mathcal{L}(\mathbf{x}_k, \boldsymbol{\lambda}_k) \begin{bmatrix} \boldsymbol{\delta}_x \\ \boldsymbol{\delta}_\lambda \end{bmatrix}$$

$$\nabla^2 \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\lambda}_k) \begin{bmatrix} \boldsymbol{\delta}_x \\ \boldsymbol{\delta}_\lambda \end{bmatrix} = -\nabla \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\lambda}_k)$$

$$\begin{bmatrix} \mathbf{W}_k & -\mathbf{A}_k^T \\ -\mathbf{A}_k & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_x \\ \boldsymbol{\delta}_\lambda \end{bmatrix} = \begin{bmatrix} \mathbf{A}_k^T \boldsymbol{\lambda}_k - \mathbf{g}_k \\ \mathbf{a}_k \end{bmatrix} \tag{15.4a}$$

Where

$$\mathbf{W}_k = \nabla_x^2 f(\mathbf{x}_k) - \sum_{i=1}^{p} (\boldsymbol{\lambda}_k)_i \nabla_x^2 a_i(\mathbf{x}_k) \tag{15.4b}$$

$$\mathbf{A}_k = \begin{bmatrix} \nabla_x^T a_1(\mathbf{x}_k) \\ \nabla_x^T a_2(\mathbf{x}_k) \\ \vdots \\ \nabla_x^T a_p(\mathbf{x}_k) \end{bmatrix} \tag{15.4c}$$

$$\mathbf{g}_k = \nabla_x f(\mathbf{x}_k) \tag{15.4d}$$

$$\mathbf{a}_k = [\, a_1(\mathbf{x}_k) \quad a_2(\mathbf{x}_k) \quad \cdots \quad a_p(\mathbf{x}_k) \,]^T \tag{15.4e}$$

$$\mathbf{W}_k \boldsymbol{\delta}_x + \mathbf{g}_k = \mathbf{A}_k^T \boldsymbol{\lambda}_{k+1} \tag{15.5a}$$

$$\mathbf{A}_k \boldsymbol{\delta}_x = -\mathbf{a}_k \tag{15.5b}$$

$$\text{minimize } \tfrac{1}{2}\boldsymbol{\delta}^T \mathbf{W}_k \boldsymbol{\delta} + \boldsymbol{\delta}^T \mathbf{g}_k \tag{15.6a}$$

$$\text{subject to: } \quad \mathbf{A}_k \boldsymbol{\delta} = -\mathbf{a}_k \tag{15.6b}$$

$$\boldsymbol{\lambda}_{k+1} = (\mathbf{A}_k \mathbf{A}_k^T)^{-1} \mathbf{A}_k (\mathbf{W}_k \boldsymbol{\delta}_x + \mathbf{g}_k) \tag{15.7}$$

Algorithm 15.1 SQP algorithm for nonlinear problems with equality constraints

Step 1

Set $\{\mathbf{x},\ \boldsymbol{\lambda}\} = \{\mathbf{x}_0,\ \boldsymbol{\lambda}_0\}$, $k = 0$, and initialize the tolerance ε.

Step 2

Evaluate $\mathbf{W}_k$, $\mathbf{A}_k$, $\mathbf{g}_k$, and $\mathbf{a}_k$ using Eqs. (15.4b) – (15.4e).

Step 3

Solve the QP problem in Eq. (15.6) for $\boldsymbol{\delta}$ and compute Lagrange multiplier $\boldsymbol{\lambda}_{k+1}$ using Eq. (15.7).

Step 4

Set $\mathbf{x}_{k+1} = \mathbf{x}_k + \boldsymbol{\delta}_x$. If $||\boldsymbol{\delta}_x|| \leq \varepsilon$, output $\mathbf{x}^* = \mathbf{x}_{k+1}$ and stop; otherwise, set $k = k + 1$, and repeat from Step 2.

SQP problems with inequality constraints

the Karush-Kuhn-Tucker (KKT) conditions

$$
\begin{aligned}
\nabla_x \mathcal{L}(\mathbf{x},\ \boldsymbol{\mu}) &= \mathbf{0} \\
c_j(\mathbf{x}) &\geq 0 \qquad \text{for } j = 1, 2, \ldots, q \\
\boldsymbol{\mu} &\geq \mathbf{0} \\
\mu_j c_j(\mathbf{x}) &= 0 \qquad \text{for } j = 1, 2, \ldots, q
\end{aligned}
$$

$$
\begin{aligned}
\nabla_x \mathcal{L}(\mathbf{x}_{k+1},\ \boldsymbol{\mu}_{k+1}) \approx\ & \nabla_x \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\mu}_k) + \nabla_x^2 \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\mu}_k)\boldsymbol{\delta}_x \\
& + \nabla_{x\mu}^2 \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\mu}_k)\boldsymbol{\delta}_\mu = \mathbf{0} \qquad (15.12\text{a})
\end{aligned}
$$

$$c_j(\mathbf{x}_k + \boldsymbol{\delta}_x) \approx c_j(\mathbf{x}_k) + \boldsymbol{\delta}_x^T \nabla_x c_j(\mathbf{x}_k) \geq 0 \text{ for } j = 1,\ 2,\ \ldots,\ q \tag{15.12b}$$

$$\boldsymbol{\mu}_{k+1} \geq \mathbf{0} \tag{15.12c}$$

and

$$[c_j(\mathbf{x}_k) + \boldsymbol{\delta}_x^T \nabla_x c_j(\mathbf{x}_k)](\boldsymbol{\mu}_{k+1})_j = 0 \qquad \text{for } \ j = 1,\ 2,\ \ldots,\ q \tag{15.12d}$$

$\downarrow$

The Lagrangian $\mathcal{L}(\mathbf{x},\ \boldsymbol{\mu})$ in this case is defined as

$$\mathcal{L}(\mathbf{x},\ \boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{j=1}^{q} \mu_j c_j(\mathbf{x}) \tag{15.13}$$

Hence

$$\nabla_x \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\mu}_k) = \nabla_x f(\mathbf{x}_k) - \sum_{j=1}^{q} (\boldsymbol{\mu}_k)_j \nabla_x c_j(\mathbf{x}_k) = \mathbf{g}_k - \mathbf{A}_k^T \boldsymbol{\mu}_k$$

$$\nabla_x^2 \mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\mu}_k) = \nabla_x^2 f(\mathbf{x}_k) - \sum_{j=1}^{q} (\boldsymbol{\mu}_k)_j \nabla_x^2 c_j(\mathbf{x}_k) = \mathbf{Y}_k \tag{15.14a}$$

and

$$\nabla^2_{x\mu}\mathcal{L}(\mathbf{x}_k,\ \boldsymbol{\mu}_k) = -\mathbf{A}_k^T$$

where $\mathbf{A}_k$ is the Jacobian of the constraints at $\mathbf{x}_k$, i.e.,

$$\mathbf{A}_k = \begin{bmatrix} \nabla_x^T c_1(\mathbf{x}_k) \\ \nabla_x^T c_2(\mathbf{x}_k) \\ \vdots \\ \nabla_x^T c_q(\mathbf{x}_k) \end{bmatrix} \tag{15.14b}$$

The approximate KKT conditions in Eq. (15.12) can now be expressed as

$$\mathbf{Y}_k\boldsymbol{\delta}_x + \mathbf{g}_k - \mathbf{A}_k^T\boldsymbol{\mu}_{k+1} = \mathbf{0} \tag{15.15a}$$

$$\mathbf{A}_k\boldsymbol{\delta}_x \geq -\mathbf{c}_k \tag{15.15b}$$

$$\boldsymbol{\mu}_{k+1} \geq \mathbf{0} \tag{15.15c}$$

$$(\boldsymbol{\mu}_{k+1})_j(\mathbf{A}_k\boldsymbol{\delta}_x + \mathbf{c}_k)_j = 0 \qquad \text{for } j = 1,\ 2,\ \ldots,\ q \tag{15.15d}$$

where

$$\mathbf{c}_k = [c_1(\mathbf{x}_k)\ c_2(\mathbf{x}_k)\ \ \ldots\ c_q(\mathbf{x}_k)]^T \tag{15.16}$$

Given $(\mathbf{x}_k,\ \boldsymbol{\mu}_k)$, Eq. (15.15) may be interpreted as the *exact* KKT conditions of the QP problem

$$\text{minimize } \tfrac{1}{2}\boldsymbol{\delta}^T\mathbf{Y}_k\boldsymbol{\delta} + \boldsymbol{\delta}^T\mathbf{g}_k \tag{15.17a}$$

$$\text{subject to:} \quad \mathbf{A}_k\boldsymbol{\delta} \geq -\mathbf{c}_k \tag{15.17b}$$

$$\downarrow$$

$$\mathbf{Y}_k\boldsymbol{\delta}_x + \mathbf{g}_k - \mathbf{A}_{ak}^T\hat{\boldsymbol{\mu}}_{k+1} = \mathbf{0}$$

$$\downarrow$$

where the rows of $\mathbf{A}_{ak}$ are those rows of $\mathbf{A}_k$ satisfying the equality $(\mathbf{A}_k\boldsymbol{\delta}_x + \mathbf{c}_k)_j = 0$ and $\hat{\boldsymbol{\mu}}_{k+1}$ denotes the associated Lagrange multiplier vector. Hence $\hat{\boldsymbol{\mu}}_{k+1}$ can be computed as

$$\hat{\boldsymbol{\mu}}_{k+1} = (\mathbf{A}_{ak}\mathbf{A}_{ak}^T)^{-1}\mathbf{A}_{ak}(\mathbf{Y}_k\boldsymbol{\delta}_x + \mathbf{g}_k) \tag{15.18}$$

Algorithm 15.2 SQP algorithm for nonlinear problems with inequality constraints

Step 1

Initialize $\{\mathbf{x},\ \boldsymbol{\mu}\} = \{\mathbf{x}_0,\ \boldsymbol{\mu}_0\}$ where $\mathbf{x}_0$ and $\boldsymbol{\mu}_0$ are chosen such that $c_j(\mathbf{x}_0) \geq 0\ (j = 1,\ 2,\ \ldots,\ q)$ and $\boldsymbol{\mu}_0 \geq \mathbf{0}$.

Set $k = 0$ and initialize tolerance ε.

Step 2

Evaluate $\mathbf{Y}_k,\ \mathbf{A}_k,\ \mathbf{g}_k$ and $\mathbf{c}_k$ using Eqs. (15.14a), (15.14b), (15.4d), and (15.16), respectively.

Step 3

Solve the QP problem in Eq. (15.17) for $\boldsymbol{\delta}_x$ and compute Lagrange multiplier $\hat{\boldsymbol{\mu}}_{k+1}$ using Eq. (15.18).

Step 4

Set $\mathbf{x}_{k+1} = \mathbf{x}_k + \boldsymbol{\delta}_x$. If $||\boldsymbol{\delta}_x|| \leq \varepsilon$, output $\mathbf{x}^* = \mathbf{x}_{k+1}$ and stop; otherwise, set $k = k + 1$, and repeat from Step 2.