Part II

constrained optimization problems

constrained optimization problems

constrained optimization algorithms

interior-point method

nonlinear constrained optimization problem

> *Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals.*

Stanislaw Ulam

constrained optimization problem: definitions

$$\min_{x\in X} f(x) \quad \text{subject to} \quad \begin{cases} c_\alpha(x) = 0, \text{ if } \alpha \in \mathcal{E} \\ c_\beta(x) \geq 0, \text{ if } \beta \in \mathcal{I}. \end{cases}$$

- A *feasible set* is a set $\Omega \subset X$ such that if $x \in \Omega$, then $c_\alpha(x) = 0$ and $c_\beta(x) \geq 0$ for $\alpha \in \mathcal{E},\ \beta \in \mathcal{I}$.
- For $x \in \Omega$ an *active set* is a subset $\mathcal{A}_x \subset \mathcal{E} \cup \mathcal{I}$ such that $\mathcal{E} \subset \mathcal{A}_x$ and $c_\beta(x) = 0$ for all $\beta \in \mathcal{A}_x$.
- For $x \in \Omega$, we say that the *linear independence constraint qualification* (LICQ) holds at x if $\partial_a c_\alpha(x)$ are linearly independent for $\alpha \in \mathcal{A}_x$.
- The *Lagrangian function* is

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{\alpha\in\mathcal{E}\cup\mathcal{I}} \lambda^\alpha c_\alpha.$$

constrained optimization problem: KKT conditions

If

1. x^* is a local minimizer, and
2. the LICQ holds at x^*,

then there exists λ_*^α such that

$$\begin{aligned} \partial_a \mathcal{L}(x^*, \lambda_*) = 0, \\ c_\alpha(x^*) = 0, \quad \forall \alpha \in \mathcal{E}, \\ c_\beta(x^*) \geq 0, \quad \forall \beta \in \mathcal{I}, \\ \lambda_*^\beta \geq 0, \quad \forall \beta \in \mathcal{I}, \\ \lambda_*^\alpha c_\alpha(x^*) = 0, \quad \forall \alpha \in \mathcal{E} \cup \mathcal{I}. \end{aligned}$$

(Karush–Kuhn–Tucker conditions)

constrained optimization problems

constrained optimization algorithms

interior-point method

merit function: l_1 penalty function

The l_1 penalty function is given by

$$\phi_1(x;\mu) = f(x) + \mu \sum_{\alpha \in \mathcal{E}} |c_\alpha(x)| + \mu \sum_{\beta \in \mathcal{I}} \max(0, -c_\beta(x)).$$

An *exact metric function* is a metric function $\phi(x;\mu)$ if there exists μ^* such that for any $\mu > \mu^*$, a local minimizer of the constrained optimization problem is a local minimizer of $\phi(x;\mu)$.

filter

The *infeasibility* is given by

$$h(x) = \sum_{\alpha \in \mathcal{E}} |c_\alpha(x)| + \sum_{\beta \in \mathcal{I}} \mathtt{max}(0, -c_\beta(x)).$$

Given a sequence of $\langle x_l : l = 1, 2, \dots, k-1 \rangle$ an iterate x_k is *acceptable* if $f(x_k) < f(x_l)$ or $h(x_k) < h(x_l)$ for $l = 1, 2, \dots, k-1$.

constrained optimization problems

constrained optimization algorithms

interior-point method

slack variables

Instead of

$$\min_{x \in X} f(x) \quad \text{subject to} \quad \begin{cases} c_\alpha(x) = 0, \ \text{if } \alpha \in \mathcal{E} \\ c_\beta(x) \geq 0, \ \text{if } \beta \in \mathcal{I}. \end{cases}$$

we can formulate the problem as

$$\min_{x \in X, s \in Y} f(x) \quad \text{subject to} \quad \begin{cases} c_\alpha(x) = 0, \ \text{if } \alpha \in \mathcal{E} \\ c_\beta(x) - s_\beta = 0, \ \text{if } \beta \in \mathcal{I} \\ s_\beta \geq 0. \end{cases}$$

KKT conditions

$$
\begin{aligned}
\partial_a f(x) - \sum_{\alpha \in \mathcal{E} \cup \mathcal{I}} \lambda^\alpha \partial_a c_\alpha(x) &= 0, \\
\sum_{\beta \in \mathcal{I}} s_\beta \lambda^\beta &= 0, \\
c_\alpha(x) = 0,\ \alpha &\in \mathcal{E}, \\
c_\beta(x) - s = 0,\ \beta &\in \mathcal{I}, \\
s_\beta \geq 0,\ \lambda_\beta \geq 0,\ \beta &\in \mathcal{I}.
\end{aligned}
$$

combinatorial complexity

$$\sum_{\beta \in \mathcal{I}} s_\beta \lambda^\beta = 0 \leftrightarrow \lambda^\beta = 0 \text{ if } \beta \notin \mathcal{A}_{x^*}$$

leads to $2^{\|\mathcal{I}\|}$ choices.

barrier method and perutrbed KKT conditions

By introducing a barrier term

$$\min_{x,s} \left(f(x) - \mu \sum_{\beta \in \mathcal{I}} \log s_\beta \right)$$

we have no combinatorial complexity.

Under some technical conditions there exists $(x(\mu), s(\mu), \lambda(\mu))$ in an open set of a solution converging to the solution.

perutrbed KKT conditions

$$
\begin{aligned}
\partial_a f(x) - \sum_{\alpha \in \mathcal{E} \cup \mathcal{I}} \lambda^\alpha \partial_a c_\alpha(x) &= 0, \\
\sum_{\beta \in \mathcal{I}} (s_\beta \lambda^\beta - \mu) &= 0, \\
c_\alpha(x) = 0,\ \alpha &\in \mathcal{E}, \\
c_\beta(x) - s = 0,\ \beta &\in \mathcal{I}, \\
s_\beta \geq 0,\ \lambda_\beta \geq 0,\ \beta &\in \mathcal{I}.
\end{aligned}
$$

interior-point method: primal-dual system

Search for a Newton direction by solving

$$\begin{pmatrix} \partial^2_{ab}\mathcal{L} & 0 & -\partial_a c_\alpha(x) & -\partial_a c_\beta(x) \\ 0 & \lambda^\beta & 0 & s_\beta \\ \partial_a c_\alpha(x) & 0 & 0 & 0 \\ \partial_a c_\beta(x) & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \delta x^a \\ \delta s^\beta \\ \delta \lambda^\alpha \\ \delta \lambda^\beta \end{pmatrix}$$

$$= \begin{pmatrix} \partial_a f(x) - \sum_{\alpha \in \mathcal{E} \cup \mathcal{I}} \partial_a c_\alpha(x) \lambda^\alpha \\ \sum_{\beta \in \mathcal{I}} (s_\beta \lambda^\beta - \mu) \\ c_\alpha(x) \\ c_\beta(x) - s \end{pmatrix}$$

interior-point method: interior-point algorithm

Algorithm 1: interior-point algorithm

Data: $x_0,\ s_0 > 0, \mu_0 > 0, \nu > 0, \tau > 0, r \in (0,1), \sigma \in (0,1)$.
Data: $\phi(x,s;\mu,\nu) = f(x) - \mu \log s + \nu \|c_\alpha(x)\| + \nu \|c_\beta(x)\|$
compute λ_0 from KKT conditions;
while $\phi(x,s;\mu,\nu) \leq \tau$ **do**
    while $\|KKT\|_{\mu_k} \leq \mu_k$ **do**
        solve the primal-dual system for $(\delta x, \delta s, \delta \lambda)$;
        $\alpha_s \leftarrow \max\{\alpha \in (0,1] : s + \alpha\delta s \geq (1-r)s\}$;
        $\alpha_z \leftarrow \max\left\{\alpha \in (0,1] : s + \alpha\delta \lambda^\beta \geq (1-r)\lambda^\beta\right\}$;
        $x_{k+1} \leftarrow x_k + \alpha_s \delta x$;
        $\lambda^\alpha_{k+1} \leftarrow \lambda^\alpha_k + \alpha_z \delta \lambda^\alpha$;
        $s_{k+1} \leftarrow s_k + \alpha_s \delta s$;
        $\lambda^\beta_{k+1} \leftarrow \lambda^\beta_k + \alpha_z \delta \lambda^\beta$;
    end
    $\mu_k \leftarrow m \in (0, \sigma\mu_k)$
end
