The Penalty Function Method

Lecture 13, Continuous Optimisation
Oxford University Computing Laboratory, HT 2006
Notes by Dr Raphael Hauser (hauser@comlab.ox.ac.uk)

I. Basic Concepts in Constrained Optimisation

In the remaining four lectures we will study algorithms for solving constrained nonlinear optimisation problems of the standard form

$$
\begin{aligned}
\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \ & f(x) \\
\text{s.t.} \quad & g_{\mathcal{E}}(x) = 0, \\
& g_{\mathcal{I}}(x) \geq 0.
\end{aligned}
$$

Two central ideas underly all of the algorithms we will consider:

- The use of *merit functions* allows one to combine the often conflicting goals of improving the objective function and achieving feasibility.

- The use of a *homotopy parameter* allows one to reduce a constrained optimisation problem to a sequence of unconstrained optimisation problems.

Merit Functions: Starting from a current iterate x, we aim at finding a new update x_+ that brings us closer towards the achievement of two conflicting goals: reducing the objective function as much as possible, and satisfying the constraints.

The two goals can be combined by minimising a *merit function* which depends both on the the objective function and on the residuals measuring the constraint violation,

$$\begin{aligned} r_{\mathcal{E}}(x) &:= g_{\mathcal{E}}(x) \\ r_{\mathcal{I}}(x) &:= (-g_{\mathcal{I}}(x))_+, \end{aligned}$$

where

$$(-g_j(x))_+ := \begin{cases} -g_j(x) & \text{if } -g_j(x) > 0, \\ 0 & \text{if } -g_j(x) \leq 0 \end{cases}$$

is the "positive part" of $-g_j$ ($j \in \mathcal{I}$).

Example 1: The *penalty function method* that will be further analysed below is based on the merit function

$$Q(x,\mu) = f(x) + \frac{1}{2\mu} \sum_{i \in \mathcal{E} \cup \mathcal{I}} \tilde{g}_i^2(x), \tag{1}$$

where $\mu > 0$ is a parameter and

$$\tilde{g}_i = \begin{cases} g_i & (i \in \mathcal{E}), \\ \min(g_i, 0) & (i \in \mathcal{I}). \end{cases}$$

Note that $Q(x,\mu)$ has continuous first but not second derivatives at points where one or several of the inequality constraints are active.

The Homotopy Idea:

The second term of the merit function forces the constraint violation to be small when $Q(x, \mu)$ is minimised over x.

We are not guaranteed that the constraints are exactly satisfied when μ is held fixed, but we can penalise constraint violation more strongly by choosing a smaller μ.

This leads to the idea of a *homotopy* or *continuation method* which is based on reducing μ dynamically and using the following idea for the outermost iterative loop:

Given a current iterate x and a value of the *homotopy parameter* μ such that x is an approximate minimiser of the *unconstrained* problem

$$\min_{y \in \mathbb{R}^n} Q(y, \mu), \tag{2}$$

reduce μ to a value $\mu_+ < \mu$ and – starting from x – apply one or several steps of an iterative algorithm for the minimisation of

$$\min_{y \in \mathbb{R}^n} Q(y, \mu_+),$$

until an approximate minimiser x_+ of this problem is reached.

Thus, the continuation approach replaces the constrained problem (NLP) by a sequence of unconstrained problems (2) for which we already studied solution methods.

II. The Penalty Function Method

Algorithm 1. (QPen)

S0 Initialisation

choose $x_0 \in \mathbb{R}^n$ % (not necessarily feasible)

choose $(\mu_k)_{\mathbb{N}_0} \searrow 0$ % (homotopy parameters)

choose $(\epsilon_k)_{\mathbb{N}_0} \searrow 0$ % (tolerance parameters)

S1 For $k = 0, 1, 2, \ldots$ repeat

$y^{[0]} := x_k$, $l := 0$

until $\|\nabla_x Q(y^{[l]}, \mu_k)\| \leq \epsilon_k$ repeat

find $y^{[l+1]}$ such that $Q(y^{[l+1]}, \mu_k) < Q(y^{[l]}, \mu_k)$

% (using an unconstrained minimisation method)

$l \leftarrow l + 1$

end

$x_{k+1} := y^{[l]}$

end.

Theorem 1: Convergence of Algorithm QPen.

- Let f and g_i be C^1 functions for all $i \in \mathcal{E} \cup \mathcal{I}$,

- let x^* be an accumulation point of the sequence of iterates $(x_k)_{\mathbb{N}_0}$ generated by Algorithm QPen, and let $(k_l)_{\mathbb{N}_0} \subseteq (k)_{\mathbb{N}_0}$ be such that $\lim_{l \to \infty} x_{k_l} = x^*$.

- Let us furthermore assume that that the set of gradients $\{\nabla g_i(x^*) : i \in \mathcal{V}(x^*)\}$ is linearly independent, where $\mathcal{V}(x^*) = \mathcal{E} \cup \{j \in \mathcal{I} : g_j(x^*) \leq 0\}$ is the index set of active, violated and equality constraints.

- For $i \in \mathcal{E} \cup \mathcal{I}$ let

$$\lambda_i^{[k]} = -\frac{\tilde{g}_i(x_{k+1})}{\mu_k}. \tag{3}$$

Then

i) x^* is feasible,

ii) the LICQ holds at x^*,

iii) the limit $\lambda^* := \lim_{l\to\infty} \lambda^{[k_l]}$ exists,

iv) (x^*, λ^*) is a KKT point.

The proof we are about to give only depends on the termination criterion in step S1 and not on the starting point $y^{[0]}$ in each iteration. We may therefore assume without loss of generality that $k_l = l$ for all $l \in \mathbb{N}_0$.

Proof:

- Using $\|\nabla_x Q(x_{k+1}, \mu_k)\| \leq \epsilon_k$ and the identity

$$\nabla_x Q(x_{k+1}, \mu_k) = \nabla f(x_{k+1}) + \frac{1}{\mu_k} \sum_{i \in \mathcal{E} \cup \mathcal{I}} \tilde{g}_i(x_{k+1}) \nabla g_i(x_{k+1}) \tag{4}$$

  in conjunction with the triangular inequality, we find

$$\Big\| \sum_{i \in \mathcal{E} \cup \mathcal{I}} \tilde{g}_i(x_{k+1}) \nabla g_i(x_{k+1}) \Big\| \leq \mu_k \Big(\epsilon_k + \|\nabla f(x_{k+1})\| \Big). \tag{5}$$

- Taking limits on the right-hand side, we find

$$\lim_{k\to\infty} \mu_k\Big(\epsilon_k + \|\nabla f(x_{k+1})\|\Big) = 0(0 + \|\nabla f(x^*)\|) = 0.$$

Therefore, the left-hand side of (5) converges to zero, and

$$\sum_{i\in\mathcal{V}(x^*)} g_i(x^*)\nabla g_i(x^*) = \sum_{i\in\mathcal{E}\cup\mathcal{I}} \tilde{g}_i(x^*)\nabla g_i(x^*) = 0.$$

- But since $\{\nabla g_i(x^*) : i \in \mathcal{V}(x^*)\}$ is linearly independent, it must be true that

$$g_i(x^*) = 0, \qquad (i \in \mathcal{V}(x^*)),$$

which shows that x^* is feasible. This settles i).

- Since x^* is feasible, we have $\mathcal{V}(x^*) = \mathcal{E} \cup \mathcal{A}(x^*)$.

- The linear independence of $\{\nabla g_i(x^*) : i \in \mathcal{V}(x^*)\}$ therefore implies that the LICQ holds at x^*, settling ii).

- Since $\epsilon_k \to 0$ and $\|\nabla_x Q(x_{k+1}, \mu_k)\| \leq \epsilon_k$, we have

$$\lim_{k\to\infty} \nabla_x Q(x_{k+1}, \mu_k) = 0.$$

  Moreover, f is continuous, so that $\lim_{k\to\infty} \nabla f(x_{k+1}) = \nabla f(x^*)$. Therefore, it follows from (4) that

$$\lim_{k\to\infty} \left(\sum_{i\in\mathcal{E}\cup\mathcal{I}} -\frac{\tilde{g}_i(x_{k+1})}{\mu_k} \nabla g_i(x_{k+1}) \right) = \nabla f(x^*). \qquad (6)$$

- Note that if $j \in \mathcal{I}$ and $g_j(x^*) > 0$ then $g_j(x_{k+1}) > 0$ and hence, $\tilde{g}_j(x_{k+1}) = 0$ for all k sufficiently large. In this case we therefore have

$$\lambda_j^* := \lim_{k\to\infty} \lambda_j^{[k]} = -\lim_{k\to\infty} \frac{\tilde{g}_i(x_{k+1})}{\mu_k} = \lim_{k\to\infty} 0 = 0. \qquad (7)$$

- On the other hand, since the LICQ holds at x^*, we have $\lim_{k\to\infty} \nabla g_i(x_k) = \nabla g_i(x^*) \neq 0$ for all $(i \in \mathcal{E} \cup \mathcal{A}(x^*))$, and hence,

$$\varphi_i^k \to \varphi_i^*, \qquad (i \in \mathcal{E} \cup \mathcal{A}(x^*)),$$

where $\varphi_i^k, \varphi_i^* : \mathbb{R}^n \to \mathbb{R}$ are the unique linear functionals such that

$$\varphi_i^k(g_j(x_k)),\ \varphi_i^*(g_j(x^*)) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

- This implies (see lecture notes for details) that

$$\lim_{k\to\infty} \lambda_i^{[k]} = -\lim_{k\to\infty} \frac{\tilde{g}_i(x_{k+1})}{\mu_k}$$
$$= \lim_{k\to\infty} \varphi_i^{k+1}\Big(\sum_{j\in\mathcal{E}\cup\mathcal{I}} -\frac{\tilde{g}_j(x_{k+1})}{\mu_k}\nabla g_j(x_{k+1})\Big) = \varphi_i^*(\nabla f(x^*)) =: \lambda_i^*$$

exists for all $(i \in \mathcal{E} \cup \mathcal{A}(x^*))$ and

$$\nabla f(x^*) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i^* \nabla g_i(x^*) = 0, \tag{8}$$

showing iii).

- The first of the KKT equations was established in (8). Moreover, we have already established that x^* is feasible and that $\lambda_j^* = 0$ for $j \in \mathcal{I} \setminus \mathcal{A}(x^*)$, showing complementarity.

- It only remains to check that $\lambda_j^* \geq 0$ for $(j \in \mathcal{A}(x^*))$.

- If $g_j(x_{k+1}) \leq 0$ occurs infinitely often, then clearly $\lambda_j \geq 0$.

- On the other hand, if $j \in \mathcal{A}(x^*)$ and $g_j(x_{k+1}) > 0$ for all k sufficiently large, then $\tilde{g}_j(x_{k+1}) = 0$ and $\lambda_j^{[k]} = 0$ for all k large, and this implies that $\lambda_i^* = 0$. $\square$

A Few Computational Issues:

It follows from the fact that the approximate Lagrange multipliers $\lambda_i^{[k]}$ converge that

$$\tilde{g}_i(x_{k+1}) = O(\mu_k)$$

for all $(i \in \mathcal{V}(x^*))$. This shows that μ_k has to be reduced to the order of precision by which we want the final result to satisfy the constraints. The Augmented Lagrangian Method which we will discuss in Lecture 14 performs much better.

The Hessian of the merit function can easily be computed as

$$
\begin{aligned}
D^2_{xx}Q(x,\mu) =& D^2 f(x) + \sum_{i\in\mathcal{E}\cup\mathcal{I}} \frac{\tilde{g}_i(x)}{\mu} D^2 g_i(x) + \frac{1}{\mu} \sum_{i\in\mathcal{V}(x)} \nabla g_i(x) \nabla g_i^{\mathsf{T}}(x) \\
&= C(x) + \frac{1}{\mu}\left(A^{\mathsf{T}}(x)A(x)\right),
\end{aligned}
$$

where $A^{\mathsf{T}}(x)$ is the matrix with columns $\{\nabla g_i(x) : i \in \mathcal{V}(x)\}$. Although $D^2_{xx}Q(x,\mu)$ is discontinuous on the boundary of the feasible domain, it can be argued that this is usually inconsequential in algorithms.

When $D^2_{xx}Q(x,\mu)$ is used for the minimisation of $Q(y,\mu_k)$ in the innermost loop of Algorithm QPen, the computations can become very ill-conditioned. For example, solving the Newton equations

$$D^2_{xx}Q(y^{[l]},\mu_k)d_l = -\nabla_x Q(y^{[l]},\mu_k) \tag{9}$$

directly can lead to large errors as the condition number of the matrix

$$C(y^{[l]}) + \frac{1}{\mu_k}\left(A^\top(y^{[l]})A(y^{[l]})\right)$$

is of order $O(\mu_k^{-1})$.

In this particular example, it is better to introduce a new dummy variable ξ_l, and to reformulate (9) as follows,

$$\begin{pmatrix} C(y^{[l]}) & A^\top(y^{[l]}) \\ A(y^{[l]}) & -\mu_k I \end{pmatrix} \begin{pmatrix} d_l \\ \xi_l \end{pmatrix} = \begin{pmatrix} -\nabla_x Q(y^{[l]}, \mu_k) \\ 0 \end{pmatrix}. \tag{10}$$

Indeed, if $(d_l, \xi_l)^\top$ satisfies (10) then d_l solves (9): $\mu_k^{-1} A d_l = \xi_l$ and $-\nabla_x Q = C d_l + A^\top \xi_l = C d_l + \mu_k^{-1} A^\top A d_l$.

The advantage of this method is that the system (10) is usually well-conditioned and the numerical results of high precision. Similar tricks can be applied when a quasi-Newton method is used instead of the Newton-Raphson method.

Reading Assignment: Lecture-Note 13.