Constrained optimization: indirect methods

Slides by
Jussi Hakanen
Post-doctoral researcher
jussi.hakanen@jyu.fi
Presented by
Mohammad Tabatabaei
Post-doctoral researcher
mohammad.tabatabaei@jyu.fi

Classification of the methods

- *Indirect methods*: the constrained problem is converted into a sequence of unconstrained problems whose solutions will approach to the solution of the constrained problem, the intermediate solutions need not to be feasible
- *Direct methods*: the constraints are taking into account explicitly, intermediate solutions are feasible

Transforming the optimization problem

- Constraints of the problem can be transformed if needed
- $g_i(x) \leq 0 \Leftrightarrow g_i(x) + y_i^2 = 0$, where y_i is a *slack variable*; constraint is active if $y_i = 0$
  - By adding y_i^2 no need to add $y_i \geq 0$
  - If $g_i(x)$ is linear, then linearity is preserved by $g_i(x) + y_i = 0, y_i \geq 0$
- $g_i(x) \geq 0 \ \Leftrightarrow -g_i(x) \leq 0$
- $h_i(x) = 0 \Leftrightarrow h_i(x) \leq 0 \ \& \ -h_i(x) \leq 0$

Examples of indirect methods

- Penalty function methods
- Barrier function methods

Penalty and barrier function methods

- Include constraints into the objective function with the help of an axulary function.
- Penalty function: penalizes for constraint violations
- Barrier function: prevents leaving the feasible region
- Resulting unconstrained problems can be solved by using the methods presented earlier in the course

Penalty function methods

- Generate a sequence of points that approach the feasible region from outside
- Constrained problem is converted into
$$\min_{x \in R^n} f(x) + r\,\alpha(x),$$
where $\alpha(x)$ is a *penalty function* and r is a *penalty parameter*
- Requirements: $\alpha(x) \geq 0 \ \forall\, x \in R^n$ and $\alpha(x) = 0$ if and only if $x \in S$

On convergence

- When $r \to \infty$, the solutions x^r of penalty function problems converge to a constrained minimizer ($x^r \to x^*$ and $r\alpha(x^r) \to 0$)
  - All the functions should be continuous
  - For each r, there should exist a solution for penalty functions problem and $\{x^r\}$ belongs to a compact subset of R^n

Examples of penalty functions

- *Can you give an example of a penalty function $\alpha(x)$?*
- For equality constraints
  - $h_i(x) = 0 \longrightarrow \alpha(x) = \sum_{i=1}^{l} (h_i(x))^2$ or
    $\alpha(x) = \sum_{i=1}^{l} |h_i(x)|^p, p \geq 2$
- For inequality constraints
  - $g_i(x) \leq 0 \longrightarrow \alpha(x) = \sum_{i=1}^{m} \max[0, g_i(x)]$ or
    $\alpha(x) = \sum_{i=1}^{m} \max[0, g_i(x)]^p, p \geq 2$

How to choose r?

- Should be large enough in order for the solutions be close enough to the feasible region
- If r is too large, there could be numerical problems in solving the penalty problems
- For large values of r, the emphasis is on finding feasible solutions and, thus, the solution can be feasible but far from optimum
- Typically r is updated iteratively
- Different parameters can be used for different constraints (e.g. $g_i \longrightarrow r_i, g_j \longrightarrow r_j$)
  - For the sake of simplicity, same parameter is used here for all the constraints

Algorithm

1) Choose the final tolerance $\epsilon > 0$ and a starting point x^1. Choose $r^1 > 0$ (*not too large*) and set $h = 1$.
2) Solve
$$\min_{x \in R^n} f(x) + r^h \alpha(x)$$
with some method for unconstrained problems (x^h as a starting point). Let the solution be $x^{h+1} = x(r^h)$.
3) Test optimality: If $r^h \alpha\left(x^{h+1}\right) < \epsilon$, stop. Solution x^{h+1} is close enough to optimum. Otherwise, set $r^{h+1} > r^h$ (e.g. $r^{h+1} = \kappa r^h$, where κ can be initialized to be e.g. 10). Set $h = h + 1$ and go to 2).

Barrier function method

- Prevents leaving the feasible region
- Suitable only for problems with inequality constraints
  - Set $\{x \mid g_i(x) < 0\ \forall i\}$ should not be empty
- Problem to be solved is

$$\min_x f(x) + r\beta(x)\ s.t.\ r \geq 0,$$

where:

- β is *a barrier function*: $\beta(x) \geq 0$ when $g_i(x) < 0\ \forall i$ and $\beta(x) \to \infty$ when x approaches boundary of S.

On convergence

- Denote $\Theta(r) = f(x^r) + r\beta(x^r)$
- Under some assumptions, the solutions x^r of barrier problems converge to a constrained minimizer ($x^r \to x^*$ and $r\beta(x^r) \to 0$) when $r \to 0^+$
  - All functions should be continuous
  - $\{x \mid g_i(x) < 0 \ \forall i\} \neq \emptyset$

Properties of barrier functions

- Nonnegative and continuous in $\{x \mid g_i(x) < 0\ \forall i\}$
- Approaches ∞ when the boundary of the feasible region is approached from inside
- Ideally: $\beta = 0$ in $\{x \mid g_i(x) < 0\ \forall i\}$ and $\beta = \infty$ in the boundary
  - Guarantees staying in the feasible region
- Examples of barrier functions
  - $\beta(x) = \sum_{i=1}^{m} -\frac{1}{g_i(x)}$
  - $\beta(x) = -\sum_{i=1}^{m} \ln\left(\min[1, -g_i(x)]\right)$

Algorithm

1) Choose the final tolerance $\epsilon > 0$ and a starting point x^1 s.t. $g_i(x) < 0\ \forall i$. Choose $r^1 > 0$, *not too small* (and a parameter $0 < \tau < 1$ for reducing r). Set $h = 1$.
2) Solve
$$\min_x f(x) + r^h \beta(x)$$
by using the starting point x^h. Let the solution be x^{h+1}.
1) Test optimality: If $r^h \beta\left(x^{h+1}\right) < \epsilon$, stop. Solution x^{h+1} is close enough to optimum. Otherwise, set $r^{h+1} < r^h$ (e.g. $r^{h+1} = \tau r^h$). Set $h = h + 1$ and go to 2).

Summary: penalty and barrier function methods

- Penalty and barrier functions are usually differentiable
- Minimum is obtained
  - Penalty function: $r^h \rightarrow \infty$
  - Barrier function: $r^h \rightarrow 0$
- Choosing the sequence r^h essential for convergence
  - If $r^h \rightarrow \infty$ or $r^h \rightarrow 0$ too slowly, a large number of unconstrained problems need to be solved
  - If $r^h \rightarrow \infty$ or $r^h \rightarrow 0$ too fast, solutions of successive unconstrained problems are far from each other and solution time increases