Indirect method

Saturday, November 11, 2017 22:01

For the glory of God

Introduction : I am expected to have a short lecture in terms of indirect method for constrained optimization to study group.

- There are two types of methods to handle constrained optimization problem { Direct method / Indirect method

- Indirect method for constrained optimization modify the problem to allow solution with unconstrained optimization techniques.
  → The influence of the constraints is captured by adding a penalty function to the objective function.

$$\phi(x, r_p) = f(x) + r_p\, p(x)$$; Constraints will be gone when the penalty function is applied.

; where $\phi(x, r_p)$ is Pseudo objective function
$f(x)$ is objective function
$p(x)$ is penalty function
r_p is Scalar multiplier ; Amplify the penalty function

why do we need to use ?

: The penalty function is necessarily steep in the region where the constraints become active.

The steepness of the penalty function may cause numerical ill-conditioning when unconstrained algorithms are applied.

⇒ To deal with the issue, r_p is successively changed during each iteration. → Even r_p is implemented, there is still a problem if r_p is increased with large numbers.

∴ This is why indirect methods are sometimes called as "Sequential Unconstrained Minimization Technique (SUMT)"

Indirect method

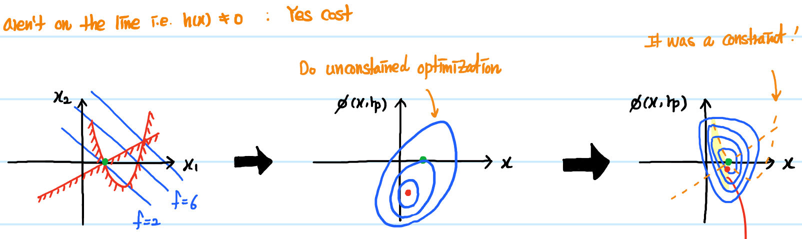
a) Exterior penalty function method

- In this method, the $p(x)$ is generally formulated as follows ;

$$p(x) = \sum_{j=1}^{m} \max[0, g_j(x)]^2 + \sum_{k=1}^{\ell} [h_k(x)]^2$$

{ If you are on the line i.e. $h(x) = 0$: No cost
  If you aren't on the line i.e. $h(x) \neq 0$: Yes cost

{ If the point is outside i.e. $g(x) > 0$: Yes cost
  If the point is inside i.e. $g(x) < 0$: No cost



- The $\phi(x, r_p)$ is then formulated as following ;

$$\phi(x, r_p) = f(x) + r_p \left\{ \sum_{j=1}^{m} \max[0, g_j(x)]^2 + \sum_{k=1}^{\ell} [h_k(x)]^2 \right\}$$

• : Minimum

$r_p = 0.05$ $r_p = 1.0$

• = minimum of $f(x_1, x_2)$
• = minimum of $\phi(x, r_p)$

Sensitive to the region where constraints used to be defined. (But close to the minimum)

It is highly recommended to start with low r_p in order to avoid numerical overflowing.

(As r_p is increased, the pseudo objective function becomes increasingly non-linear but not discontinuity for this case)
↳ when unconstrained optimization is performed w.r.t. $\phi(x, r_p)$.

- Any points in the domain could be an initial point but { If the point is in feasible region, $\phi(x, r_p)$ will be same regardless of r_p.
  If the point is in infeasible region, ϕ will be different as r_p changes.

· Therefore, even though this method is simple to implement, the result would be in infeasible region if the iteration is stopped prematurely.

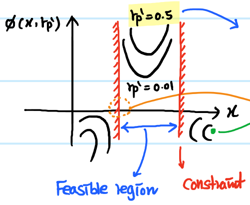
b) Interior penalty function method

· Since we presume that the initial search is begun in the feasible region, a feasible solution would be obtained even if it stops early.

· However, the ø function for this case is discontinuous at the constraint boundaries, which may cause considerable problems for line search method if the initial point is defined in infeasible region.

$$\phi(x, r_p') = f(x) + r_p' \sum_{j=1}^{m} \frac{-1}{g_j(x)}$$

; Here, interior penalty function is only applied for inequality constraints because there is no interior of feasible for equality constraints.



This method is also dependent on scalar multiplier. (rp' should be decreased to obtain optimum solution)

If you start from here, there is no way to bring it into feasible region.

when $g_j(x) = 0$, then $p(x) \to \infty$ (Discontinuity)

↳ Hence, rp' is initialized as a large number to decrease the sensitivity.

c) Augmented Lagrangian Method

· In exterior penalty method, a scalar multiplier (rp) was introduced to prevent ø function from numerical ill-condition at constraint boundaries.

↳ However, as rp is increased, the ø function was increasingly non-linear when unconstrained optimization is performed.

· In interior penalty method, even though the ø function wasn't supposed to be non-linear by nature, the ø function is discontinuous at constraint boundaries, which results in that there is no way for initial points to get into the feasible region if the initial points start from infeasible.

↳ rp' was introduced to decrease the sensitivity.

· The Augmented Lagrangian method is considered to be a more modern form of penalty function method because it's relatively insensitive to rp.

- For example,

It can be started from either feasible or infeasible

· Let's compare the Lagrangian and the Pseudo-objective function with Exterior penalty method.

why not Interior ? This interior method is only valid with feasible region.

$$\text{Lagrangian function } (\mathcal{L}) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{k=1}^{\ell} \lambda_{m+k} h_k(x)$$

$$\text{Exterior penalty function } (\phi) = f(x) + r_p \left\{ \sum_{j=1}^{m} \max[0, g_j(x)]^2 + \sum_{k=1}^{\ell} [h_k(x)]^2 \right\}$$

· If we assume that x^* is an optimum for both cases, then it is true that $\begin{cases} \nabla \phi(x^*, r_p) = 0 \\ \nabla \mathcal{L}(x^*, \lambda) = 0 \end{cases}$ under the KKT condition

· Then, we may be able to say ;

$\nabla \mathcal{L}(x^*, \lambda) = \nabla \phi(x^*, \lambda)$ at optimum point → This is because $\begin{cases} f(x) = x^2 - 2x \to \nabla f(1) = 2x - 2 = 0 \\ g(x) = \frac{3}{2}x^2 - 3x \to \nabla g(1) = 3x - 3 = 0 \end{cases}$

$\therefore \nabla f(1) = \nabla g(1)$ even if $\nabla f(x) \neq \nabla g(x)$

· Let's take the gradient and compare each term.

$$\nabla \mathcal{L}(x^*, \lambda) = \nabla f(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) + \sum_{k=1}^{\ell} \lambda_{m+k} \nabla h_k(x^*)$$

$$\nabla \phi(x^*, r_p) = \nabla f(x^*) + r_p \left\{ \sum_{j=1}^{m} 2 g_j(x^*) \nabla g_j(x^*) + \sum_{k=1}^{\ell} 2 h_k(x^*) \nabla h_k(x^*) \right\}$$

Therefore, we have

$$\lambda_j = 2 r_p g_j(x^*) \quad \text{and} \quad \lambda_{m+k} = 2 r_p h_k(x^*)$$

$$\Leftrightarrow g_j(x^*) = \frac{\lambda_j}{2 r_p} \quad \text{and} \quad h_k(x^*) = \frac{\lambda_{m+k}}{2 r_p}$$

· If we think about the case in which the $g_j(x)$ is active ($g_j(x) = 0$), then r_p must be infinite.

(or for Equality constraint, $h_k(x) = 0$)

Let's compare it with Augmented Lagrangian method.

↳ First of all, we will talk about the Augmented Lagrangian method with only "Equality constraint."

· The Augmented Lagrangian method with only Equality constraints can be defined as following;

$$A(x, \lambda_p, r_p) = \mathcal{L}(x, \lambda_p) + r_p \sum_{k=1}^{\ell} [h_k(x)]^2$$

$$= f(x) + \sum_{k=1}^{\ell} \lambda_{p,k} h_k(x) + r_p \sum_{k=1}^{\ell} [h_k(x)]^2$$

· Then, $\nabla A = \nabla f(x^*) + \sum_{k=1}^{\ell} \left\{ \lambda_{p,k} \nabla h_k(x^*) + 2 r_p h_k(x^*) \nabla h_k(x^*) \right\}$

· In the same way, Compare $\nabla A(x^*, \lambda_p, r_p)$ with $\nabla \mathcal{L}(x^*, \lambda)$ because we just introduce $A(x, \lambda_p, r_p)$ instead of $\phi(x, r_p)$ as a new method.

Keep in mind that we only concern about Equality constraints case.

$$\therefore \lambda_{m+k} = \lambda_{p,k} + 2 r_p h_k(x^*)$$

$$\Leftrightarrow h_k(x^*) = \frac{\lambda_{m+k} - \lambda_{p,k}}{2 r_p}$$ → In order to obtain $h_k(x) = 0$, $r_p \to \infty$ is not only answer but we could $\lambda_{p,k} \to \lambda_{m+k}$.

(It's relatively insensitive to r_p)