International Journal of Mathematics And Its Applications

Vol.2 No.3 (2014), pp.59-70.

ISSN: 2347-1557(online)

Determination of Feasible Directions by Successive Quadratic Programming and Zoutendijk Algorithms: A Comparative Study

Tripti Sharma[†1] **and Semeneh Hunachew**[‡]

[†] *Addis Ababa Science and Technology University, Ethiopia.*

[‡] *Arba Minch University, Ethiopia.*

Abstract : This study is focused on comparison between Method of Zoutendijk and Successive Quadratic Programming (SQP) Method, which are the methods to find feasible directions while solving a non-linear programming problem by moving from a feasible point to an improved feasible point.

Keywords : Zoutendijk Algorithm, Successive Quadratic Programming Method, optimization, Feasible points, Feasible directions.

1 Introduction

ZOUTENDIJK ALGORITHM

In the Zoutendijk method of finding feasible directions, at each iteration, the method generates an improving feasible direction and then optimizes along that direction.

Definition 1.1. *Consider the problem to minimize $f(x)$ subject to $x \in S$, where $f : R^n \to R$ and S is a non empty set in R^n. A non zero vector d is called a feasible direction at $x \in S$ if there exists a $\delta > 0$ such that $x + \lambda d \in S$ for all $\lambda \in (0, \delta)$. Furthermore, d is called an improving feasible direction at $x \in S$ if there exists a $\delta > 0$ such that $f(x + \lambda d) < f(x)$ and $x + \lambda d \in S$ for all $\in (0, \delta)$. In Case of linear constraints, first consider the case where the*

[1] Corresponding author E-Mail: drtripti2010@gmail.com (Tripti Sharma)

feasible region S is defined by a system of linear constraints, so that the problem under consideration is of then form:

$$\text{Minimize } f(x)$$
$$\text{Subject to: } Ax \le b$$
$$Qx = q$$

Where, $A = m \times n$ matrix, $Q = l \times n$ matrix, $b = m-$matrix, $q = l-$matrix.

Lemma 1.2. *Consider the problem to minimize $f(x)$ subject to $Ax \le b$ and $Qx = q$. Let x be a feasible solution and suppose that $A_1x = b_1$ and $A_2x < b_2$, where A^T is decomposed in to (A_1^T, A_2^T) and b^T is decomposed in to (b_1^T, b_2^T). Then a non zero vector d is a feasible direction at x if and only if $A_1d \le 0$ and $Qd = 0$. If $\nabla f(x)^T d < 0$, then d is an improving direction.*

Generating Improving Feasible Directions

Given a feasible point x as shown in lemma 1.2, a non zero vector d is an improving feasible direction if $\nabla f(x)^T d < 0$, $A_1d \le 0$ and $Qd = 0$. A natural method for generating such a direction is to minimize $\nabla f(x)^T d$ subject to the constraints $A_1d \le 0$ and $Qd = 0$. Note, however, that if a vector $\bar{d}$ such that $\nabla f(x)^T \bar{d} < 0$, $A_1\bar{d} \le 0$, $Q\bar{d} = 0$ exists, then the optimal objective value of the forgoing problem is $-\infty$ by considering $\lambda\bar{d}$, where $\lambda \to \infty$. Thus a constraint that bounds the vector d or the objective function must be introduced. Such a restriction is usually referred to as a normalization constraint. There are the following three problems for generating an improving feasible direction. Each of the problems uses a different normalization constraint.

Problem (P1):
Minimize $\nabla f(x)^T d$
Subject to: $A_1x \le b$
$Qx = q; \quad -1 \le d_j \le 1$ for $j = 1, 2, ..., n$

Problem (P2):
Minimize $\nabla f(x)^T d$
Subject to: $A_1x \le b$
$Qx = q, \quad d^T d \le 1$

Problem (P3):
Minimize $\nabla f(x)^T d$
Subject to: $A_1x \le b$
$Qx = q; \quad \nabla f(x)^T d \ge -1$

Problem P_1 and P_2 are linear in the variables $d_1, ..., d_n$ and can be solved by the simplex method. Problem P_2 contains a quadratic constraint but could be considerably simplified. Since $d = 0$ is a feasible solution to each of the above problems and since its objective value is equal to zero, the optimal objective value of problems P_1, P_2 and P_3 cannot be positive. If the minimal objective function value of P_1, P_2 and P_3 is negative then by lemma 1.2; an improving g feasible direction is generated. On the other hand, if the minimal objective function value is equal to zero, then x is a KKT point as shown below.

Lemma 1.3. *Consider the problem to minimize $f(x)$ subject to $Ax \le b$ and $Qx = q$. Let x be a feasible solution such that $A_1x = b_1$ and $A_2x < b_2$, where $AT = (A_1^T, A_2^T)$ and $b^T = (b_1^T, b_2^T)$. Then for each $i = 1, 2, 3$, x is a KKT point if and only if the optimal objective value of problem P_i is equal to zero.*

Line Search: Let x_k be the current vector, and let d_k be an improving feasible direction. The next x_{k+1} is given by $x_k + \lambda_k d_k$, where the step size λ_k is obtained by solving the following one dimensional problem:

Minimize $f(x_k + \lambda_k d_k)$

Subject to: $A(x_k + \lambda_k d_k) \le b$

$Q(x_k + \lambda_k d_k) = q; \quad \lambda \ge 0$

Now, suppose that A^T is decomposed in to (A_1^T, A_2^T) and b^T is decomposed into (b_1^T, b_2^T) such that $A_1x_k = b_1$ and $A_2x_k \le b2$. Then the above problem could be simplified as follows:

First note that $Qx_k = q$ and $Qx_k = 0$, So that the constraint $Q(x_k + \lambda_k d_k) = q$ is redundant. Since $A_1x_k = b_1$ and $A_1d_k \le 0$, then $A_1(x_k + \lambda_k d_k) \le b_1$ for all $\lambda \ge 0$. Hence, we only need to restrict λ so that $\lambda A_2 d_k \le b2 - A_2 x_k$ and the above problem reduce to the following line search problem i.e;

$$\begin{aligned}
&\text{Minimize } \; f(x_k + \lambda_k d_k)\\
&\text{Subject to: } \; 0 \le \lambda \le \lambda_{\max},\\
&\text{where } \quad \lambda_{\max} = \begin{cases} \min\left\{\frac{\tilde{b_1}}{\tilde{d_1}} : \; \tilde{d_1} > 0\right\}, & \text{if } \tilde{d} > 0 \\ \infty, & \text{if } \tilde{d} \le 0 \end{cases}\\
&\tilde{b} = b_2 - A_2x_k \; \text{ and } \; \tilde{d} = A_2x_k
\end{aligned} \tag{1.1}$$

In case of Linear Constraints

Consider the problem (P):

Minimize $f(x)$

Subject to: $\quad Ax \le b; \quad Qx = q$

Initial Step: Find a starting feasible solution x_1 with $Ax_1 \le b$ and $Qx_1 = q$. Let $k = 1$ and go to the main step:

Main Step:

1. Given x_k, suppose that A^T and b^T are decomposed in to $(A_1^T, A_2^T)(b_1^T, b_2^T)$ So that $A_1x_k = b_1$ and $A_2x_k \le b2$. Let d_k be an optimal solution to the following problem(note that problem p_2 or p_3 could be used instead):

$$\begin{aligned}
&\text{Minimize } \; \nabla f(x)^T d\\
&\text{Subject to: } \; A_1 d \le 0\\
&\qquad\qquad Qd = 0\\
&\qquad -1 \le d_j \le 1 \; \text{ for } \; j = 1, ..., n
\end{aligned}$$

If $\nabla f(x)^T d_k = 0$, Stop; x_k is KKT point, with the dual variables to the forgoing problem giving the corresponding Lagrange multipliers. Let λ_k be an optimal solution to the following line search problem: Otherwise, go to step 2

Let λ_k be an optimal solution to the following line search problem:

Minimize $f(x_k + \lambda d_k)$

subject to: $0 \leq \lambda \leq \lambda_{\max}$

Where $\lambda_{\max}$ is determined according to (2.1a). Let $x_{k+1} = x_k + \lambda_k d_k$. Identify the new set of binding constraints at x_{k+1}, and update A_1 and A_2 accordingly. Replace k by $k+1$ and go to step 1.

Problems with non-linear inequality constraints

The following theorem gives the sufficient condition for a vector d to be an improving feasible direction.

Theorem 1.4. *Consider the following problem*

Minimize $f(x)$

subject to: $g_i(x) \leq 0$ *for* $i = 1, ..., m$.

Let x be a feasible solution, and let I be the set of binding or active constraints, that is $I = \{i : g_i(x) = 0\}$. Furthermore, suppose that f and g_i for $i \in I$ are differentiable at x and that each g_i for $i \notin I$ is continuous at x. If $\nabla f(x)^T d < 0$ and $\nabla g_i(x)^T d < 0$ for $i \in I$, then d is an improving direction.

Theorem 1.5. *Consider the problem to minimize $f(x)$ subject to $g_i(x) \leq 0$ for $i = 1, 2, 3, ..., m$. Let x be a feasible solution and let $I = \{i : g_i(x) = 0\}$. Consider the following direction finding problem:*

$$\begin{aligned} &\textit{Minimize } z \\ &\textit{Subject to: } \nabla f(x)^T d - z \leq 0 \\ &\qquad \nabla g_i(x)^T d - z \leq 0 \;\; \textit{for } i \in I \\ &\qquad -1 \leq d_j \leq 1 \;\; \textit{for } j = 1, ..., n \end{aligned}$$

Then x is a Fritz John point if and only if the optimal objective value to the above problem is zero.

In Case of Non-linear Inequality Constraints

Initial step: Choose a starting point x_1 such that $g_i(x_1) \leq 0$ for $i = 1, 2, 3, ..., m$. Let $k = 1$ and go to the main step.

Main step:

1. $I = \{i : g_i(x_k) = 0\}$ and solve the following problem:

$$\begin{aligned} &\text{Minimize } z \\ &\text{Subject to: } \nabla f(x_k)^T d - z \leq 0 \\ &\qquad \nabla g_i(x_k)^T d - z \leq 0 \;\; \text{for } i \in I \\ &\qquad -1 \leq d_j \leq 1 \;\; \text{for } j = 1, ..., n \end{aligned}$$

Let (z_k, d_k) be an optimal solution. If $z_k = 0$, stop; x_k is a fritz John point. If $z_k < 0$, then go to step 2.

2. Let λ_k be an optimal solution to the following line search problem:

$$\text{Minimize } f(x_k + \lambda d_k)$$
$$\text{Subject to: } 0 \le \lambda \le \lambda_{\max}$$

Where, $\lambda_{\max} = \sup\{\lambda : g_i(x_k + \lambda d_k) \le 0 \text{ for } i = 1, 2, 3, ..., m\}$. Let $x_{k+1} = x_k + \lambda_k d_k$, replace k by $k+1$ and go to step 1.

2 Topkis-Veintt's modification of the feasible direction algorithm

A modification of Zoutendijks method of feasible directions was proposed by Topkis and Veinott [1967] and guarantees convergence to a Fritz John point. The problem under consideration is given by

$$\text{Minimize } f(x)$$
$$\text{Subject to: } g_i(x) \le 0 \text{ for } i = 1, ..., m$$

Generating a Feasible Direction

Given a feasible point x, a direction is found by solving the following direction finding linear programming problem DF(x):

$$\text{Problem DF(x):} \quad \text{Minimize z}$$
$$\text{Subject to: } \nabla f(x)^T d - z \le 0$$
$$\nabla g_i(x)^T d - z \le -g_i(x) \text{ for } i = 1, ..., m$$
$$-1 \le d_j \le 1 \text{ for } j = 1, ..., n$$

Here both binding and non binding constraints play a role in determining the direction of movement. As opposed to the method of feasible direction of approaching the boundary of a currently nonbinding constraint.

Topkis-Veinotts Algorithm

Initial Step:

Choose a point x_1 such that $g_i(x_1) \le 0$ for $i = 1, ..., m$. Let $k = 1$ and go to the main step

Main step:

1. Let (z_k, d_k) be an optimal solution to the following linear programming problem :

$$\text{Minimize z}$$
$$\text{Subject to: } \nabla f(x_k)^T d - z \le 0$$
$$\nabla g_i(x_k)^T d - z \le -g_i(x_k) \text{ for } i = 1, ..., m$$
$$-1 \le d_j \le 1 \text{ for } j = 1, ..., n$$

If $z_k = 0$, stop: x_k is a Fritz John point. Otherwise, $z_k \le 0$ and go to step 2.

2. Let λ_k be an optimal solution to the following line search problem:

$$\text{Minimize} \quad f(x_k + \lambda d_k)$$
$$\text{Subject to: } 0 \leq \lambda \leq \lambda_{\max},$$

Here $\lambda_{\max} = \sup\{\lambda : \; g_i(x_k + \lambda d_k) \leq 0 \text{ for } i = 1, ..., m\}$. Let $x_{k+1} = x_k + \lambda_k d_k$, replace k by $k+1$ and go to step 1.

Theorem 2.1. *Let x be a feasible solution to the problem to minimize $f(x)$ subject to $g_i(x) \leq 0$ for $i = 1, ..., m$. Let $(\bar{z}, \bar{d})$ be an optimal solution to the problem $DF(x)$. If $\bar{z} \leq 0$, then $\bar{d}$ is an improving feasible direction. Also, $\bar{z} = 0$ if and only if x is a Fritz John point.*

Lemma 2.2. *Let s be a non empty set in R^n and let $f : R^n \to R$ be continuously differentiable. Consider the problem to minimize $f(x)$ subject to $x \in S$. Further more, consider any feasible direction algorithm whose map $A = MD$ is defined as follows . Given $x, (x, d) \in D(x)$ means that d is an improving feasible direction of f at x. Furthermore, $y \in M(x, d)$ means that $y = x + \bar{\lambda} d$, where $\bar{\lambda}$ solves the line search problem to minimize $f(x + \lambda d)$ subject to $\lambda \geq 0$ and $x + \lambda d \in S$. Let $\{x_k\}$ be any sequence generated by such an algorithm, and let $\{d_k\}$ be the corresponding sequence of directions. Then there cannot exist a subsequence $\{(x_k, d_k)\}_{\mathcal{K}}$ satisfying the following properties:*

i) $x_k \to x$ *for* $k \in \mathcal{K}$

ii) $d_k \to d$ *for* $k \in \mathcal{K}$

iii) $x_k + \lambda d_k \in S$ *for all* $\lambda \in [0, \delta]$ *and for each* $k \in \mathcal{K}$ *for some* $\delta > 0$

iv) $\nabla f(x)^T d < 0$

Theorem 2.3. *Let $f, \; g_i : R^n \to R$ for $i = 1, ..., m$ be continuously differentiable, and consider problem to minimize $f(x)$ subject to $g_i(x) \leq 0$ for $i = 1, ..., m$. Suppose that the sequence $\{x_k\}$ is generated by the algorithm of Topkis and Veinott. Then any accumulation point of $\{x_k\}$ is a Fritz John point.*

3 Successive Quadratic Programming (SQP) Algorithm

SQP methods, also known as sequential, or recursive, quadratic programming approaches, employ Newtons method (or quasi-Newton methods) to directly solve the KKT conditions for the original problem. As a result, the accompanying sub-problem turns out to be the minimization of a quadratic approximation to the Lagrange function optimized over a linear approximation to the constraints. Hence, this type of process is also known as a projected Lagrangian, or Lagrange-Newton approach. By its nature this method produces both primal and dual(Lagrange multiplier) solutions.

SQP for equality constrained problem

Consider the following equality constrained problem (P):

P: Minimize f(x) (ECP)

subject to: $h(x) = 0, \, x \in R^n$

Where $f : R^n \to R$ and $h : R^n \to R^m$ are assumed to be continuously twice differentiable (smooth) functions. An understanding of this problem is essential in the design of SQP methods for general non-linear programming problems. The KKT optimality conditions for problem P require a primal solution $x \in R^n$ and a Lagrange multiplier vector $\lambda \in R^m$ such that

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) = 0 \qquad h_i(x) = 0, \quad i = 1, ..., m \tag{3.1}$$

If we use the Lagrangian

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) \tag{3.2}$$

we can write the KKT conditions (4.1a) more compactly as $\begin{pmatrix} \nabla_x L(x,\lambda) \\ \nabla_\lambda L(x,\lambda) \end{pmatrix} = 0$ (EQKKT) The main idea behind SQP is to model problem (ECP) at the given point $x^{(k)}$ by a quadratic programming subproblem and then use the solution to this problem to construct a more accurate approximation $x^{(k+1)}$.If we perform a Taylor series expansion of the system (EQKKT) about $(x^{(k)}, \lambda^{(k)})$ we obtain

$$\begin{pmatrix} \nabla_x L(x^{(k)}, \lambda^{(k)}) \\ \nabla_\lambda L(x^{(k)}, \lambda^{(k)}) \end{pmatrix} + \begin{pmatrix} \nabla_x^2 L(x^{(k)}, \lambda^{(k)}) & \nabla h(x^{(k)}) \\ \nabla h(x^{(k)})^T & 0 \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = 0$$

Where $\delta_x = x^{(k+1)} - x^{(k)}$, $\delta_\lambda = \lambda^{(k+1)} - \lambda^{(k)}$ and $\nabla_x^2 L(x, \lambda) = \nabla^2 f(x) + \sum\limits_{i=1}^{m} \lambda_i \nabla^2 h_i(x)$ is the Hessian matrix of the Lagrangian function. Taylor series expansion can be written equivalently as

$$\begin{pmatrix} \nabla_x^2 L(x^{(k)}, \lambda^{(k)}) & \nabla h(x^{(k)}) \\ \nabla h(x^{(k)})^T & 0 \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = \begin{pmatrix} -\nabla f(x^{(k)}) - \nabla h(x^{(k)}) \lambda^{(k)} \\ -h(x^{(k)}) \end{pmatrix}$$

or, setting $d = \delta_x$ and bearing in mind that

$$\lambda^{(k+1)} = \delta_\lambda + \lambda^{(k)} \begin{pmatrix} \nabla_x^2 L(x^{(k)}, \lambda^{(k)}) & \nabla h(x^{(k)}) \\ \nabla h(x^{(k)})^T & 0 \end{pmatrix} \begin{pmatrix} d \\ \lambda^{(k+1)} \end{pmatrix} = \begin{pmatrix} -\nabla f(x^{(k)}) \\ -h(x^{(k)}) \end{pmatrix} \tag{3.3}$$

Algorithm of SQP method

1. Determine $(x^{(0)}, \lambda^{(0)})$
2. Set $K = 0$
3. Repeat until convergence test is satisfied
4. Solve the system (3.2) to determine $(d^{(k)}, \lambda^{(k+1)})$
5. Set $x^{(k+1)} = x^{(k)} + d^{(k)}$
6. Set $k = k + 1$
7. End (got to step 2)

In SQP methods, problem (ECP) is modelled by a quadratic programming sub-problem (QPS for short), whose optimality conditions are the same as in the system (4.1c). The algorithm is also the same as that of SQP method, but instead of solving the system (4.1c) in step 4, we solve the following quadratic programming sub-problem (QPS):

$$\begin{aligned} &\text{Minimize} \quad \nabla f(x^{(k)})^T d + \frac{1}{2} d^T \nabla_x^2 L(x^{(k)}, \lambda^{(k)}) d \\ &\text{Subject to:} \quad h(x^{(k)}) + \nabla h(x^{(k)})^T d = 0 \end{aligned} \tag{3.4}$$

Since the first order conditions for the previous problem at $(x^{(k)}, \lambda^{(k)})$ are given by the system (3.3) and therefore $d^{(k)}$ is a stationary point of (3.4). If $d^{(k)}$ satisfies second order sufficient conditions, then $d^{(k)}$ minimizes problem (3.4). We also observe that the constraints in (3.4) are derived by a first order Taylor series approximation of the constraints of the original problem (ECP). The objective function of the QPS is a truncated second order Taylor series expansion of the Lagrangian function.

Convergence Rate Analysis

Under appropriate conditions, we can argue a quadratic convergence behavior for the or going algorithm. Specifically, suppose that $\bar{x}$ is a regular KKT solution for problem p which together with a set of Lagrange multipliers, $\bar{x}$ satisfies the second order sufficient conditions. Then $\nabla W(\bar{x}, \bar{\lambda}) = \begin{pmatrix} \nabla_x^2 L(\bar{x}, \bar{\lambda}) & \nabla h(\bar{x}) \\ \nabla h(\bar{x})^T & 0 \end{pmatrix}$ is non-singular. To see this, let us show that the system $\nabla W(\bar{x}, \bar{\lambda}) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 0$ has the unique solution given by $(d_1^T, d_2^T) = 0$. Consider any solution (d_1^T, d_2^T). Since $\bar{x}$ is a regular solution, $\nabla h(\bar{x})^T$ has full rank; so if $d_1 = 0$, then $d_2 = 0$ as well.

If $d_1 \neq 0$, Since $\nabla h(\bar{x})^T d_1 = 0$, we have by the second order sufficient conditions that $d_1^T \nabla^2 L(\bar{x}) d_1 > 0$. However since $\nabla^2 L(\bar{x}) d_1 + \nabla h(\bar{x})^T d_2 = 0$, we have that $\nabla^2 L(\bar{x}) d_1 = -d_2^T \nabla h(\bar{x}) d_1 = 0$, a contradiction. Hence $\nabla W(\bar{x}, \bar{\lambda})$ is non-singular and thus for (x_k, λ_k) sufficiently close to $(\bar{x}, \bar{\lambda})$, $\nabla W(x_k, \lambda_k)$ is non-singular.

Extension to Include Inequality Constraints

The sequential quadratic programming framework can be extended to general non-linear constrained problem

$$\begin{aligned} &\text{P: Minimize} \quad f(x) \\ &\text{Subject to:} \quad h_i(x) = 0, \quad i = 1, ..., m \\ &\qquad\qquad\qquad g_i(x) \leq 0, \quad i = 1, ..., l \end{aligned} \tag{3.5}$$

Where, f, h_i, g_i are continuously twice differentiable for each i. For instance, given an iterative (x_k, u_k, v_k) where $u_k \geq 0$ and v_k are respectively, the Lagrange multiplier estimates for the inequality and the equality constrains, we consider the following quadratic sub-problem.

$$\begin{aligned} &\text{Minimize } \ f(x_k) + \nabla f(x_k)^T d + \frac{1}{2} d^T \nabla_x^2 L(x_k, \lambda_k) d \\ &\text{Subject to:} \\ &\qquad h_i(x_k) + \nabla h_i(x_k)^T d = 0, \quad i = 1, ..., m \\ &\qquad g_i(x_k) + \nabla g_i(x_k)^T d \leq 0, \quad i = 1, ..., l \end{aligned} \tag{3.6}$$

Where $\nabla_x^2 L(x_k, \lambda_k) = \nabla^2 f(x_k) + \sum_{i=1}^{l} u_{k_i} \nabla^2 g_i(x_k) + \sum_{i=1}^{m} v_{k_i} \nabla^2 h_i(x_k)$. Note that the KKT conditions for this problem require that in addition to primal feasibility, we find Lagrange multipliers u and v such that

$$\nabla f(x_k) + \nabla_x^2 L(x_k, \lambda_k) d + \sum_{i=1}^{l} u_i \nabla g_i(x_k) + \sum_{i=1}^{m} v_i \nabla h_i(x_k) = 0$$

$$u_i [g_i(x_k) + \nabla g_i(x_k) T d] = 0, \quad i = 1, .., l \quad u \geq 0, \text{ v unrestricted.} \tag{3.7}$$

Hence, if d_k solves (3.5) with Lagrange multipliers u_{k+1} and v_{k+1} and if $d_k = 0$, then x_k along with (u_{k+1}, v_{k+1}) yields a KKT solution for the original problem (P). Otherwise, we set $x_{k+1} = x_k + d_k$ as before, increment k by 1, and repeat the process .In similar manner, It can be shown that if $\bar{x}$ is a regular KKT solution which, together with $(\bar{u}, \bar{v})$ satisfies the second order sufficiency conditions, and if (x_k, u_k, v_k) is initialized sufficiently close to $(\bar{x}, \bar{u}, \bar{v})$, the forgoing iterative process will converge quadratically to close $(\bar{x}, \bar{u}, \bar{v})$.

Lemma 3.1. *Given an iterate x_k, consider the quadratic subproblem QP given by (3.6) where $\nabla_x^2 L(x_k, \lambda_k)$ is replaced by any positive definite approximation B_k. Let d solve this problem with Lagrange multipliers u and v associated with the inequality and the equality constraints, respectively. If $d \neq 0$, and if $\mu \geq \max\{u_1, ..., u_l, |v_1|, ..., |v_m|\}$, then d is a descent direction at $x = x_k$ for the merit function F_E given above.*

4 Summary

Zoutendijks Feasible Directions

- Basic idea:
  - ✓ move along steepest descent direction until constraints are encountered
  - ✓ at constraint surface, solve sub-problem to find descending feasible direction
  - ✓ repeat until KKT point is found
- Method
  - ✓ Sub-problem linear: efficiently solved
  - ✓ Determine active set before solving sub-problem!
  - ✓ When $a = 0$: KKT point found
  - ✓ Method needs feasible starting point.
- Convergence

✓ The direction-finding problem only uses the binding constraints.

✓ Nearly binding constraints can cause very short steps to be taken and also drastic changes in direction.

✓ This causes the algorithmic map not to be closed.

✓ This can cause jamming and slow convergence.

✓ Idea: Use constraints that are nearly binding in the direction-finding problem.

✓ Even this is not enough to guarantee convergence

Method of Topkis and Veinott

- Try to eliminate drastic changes in direction by accounting for all constraints.
- Use the following direction-finding problem:

Minimize z

Subject to:

$$\nabla f(x^*)^T d - z \leq 0$$

$$\nabla g_i(x^*)^T d - z \leq -g_i(x^*), \quad -1 \leq d_j \leq 1$$

This is enough to guarantee convergence to an FJ point.

- Convergence of Topkis and Veinott

✓ Note that the solution to the direction-finding problem is feasible and improving.

✓ Also, the optimal solution is 0 if and only if the current point is an FJ point.

✓ Taking all the constraints into account eliminates drastic changes in direction and ensures that the algorithmic map is closed.

✓ Under the assumption that all the functions involved are continuously differentiable, a sequence $\{x_k\}$ is generated by this algorithm, then allaccumulation points are FJ points.

Method of Topkis and Veinott

- Basic idea

✓ The basic idea is analogous to Newtons method for unconstrained optimization.

✓ In unconstrained optimization, only the objective function must be approximated, in the NLP, both the objective and the constraint must be modeled.

✓ An sequential quadratic programming method uses a quadratic for the objective and a linear model of the constraint (i.e., a quadratic program at each iteration)

✓ Solve the KKT conditions directly using a Newton method.

✓ This leads to a method which amounts to minimizing a second-order approximation of the Lagrangian.

✓ From this, we can get a quadratic convergence rate

$$\text{Minimize } f(x)$$
$$\text{Subject to:}$$
$$h_i(x) = 0, \quad i = 1, ..., m$$
$$g_i(x) \le 0, \quad i = 1, ..., l$$
$$\Downarrow$$

$$\text{Minimize } f(x_k) + \nabla f(x_k)^T d + \frac{1}{2} d^T \nabla_x^2 L(x_k, \lambda_k) d$$
$$\text{Subject to:}$$
$$h_i(x_k) + \nabla h_i(x_k)^T d = 0, \quad i = 1, ..., m \qquad \Longrightarrow \quad x^{(k+1)} = x^{(k)} + d^{(k)}$$
$$g_i(x_k) + \nabla g_i(x_k)^T d \le 0, \quad i = 1, ..., l$$

Basic SQP Algorithm

1. Choose initial point x_0 and initial multiplier estimates λ_0
2. Set up matrices for QP sub-problem
3. Solve QP sub-problem $\rightarrow d_k, \lambda_{k+1}$
4. Set $x_{k+1} = x_k + d_k$
5. Check convergence criteria $\longrightarrow$ Finished. Otherwise go to 2.

5 Conclusion

Comparison of Zounendijk and SQP method :

S.No	Criteria	Zoutendijk	SQP
1	Feasible starting point?	Yes	No
2	Nonlinear constraints?	Yes	Yes
3	Equality constraints?	Hard	Yes
4	Uses active set?	Yes	Yes
5	Iterates feasible?	Yes	No
6	Derivatives needed?	Yes	Yes

Table: Comparison of Zounendijk and SQP method

Generally, SQP seen as best general-purpose method for constrained problems. It relies on a profound theoretical foundation and provides powerful algorithmic tools for the solution of large-scale technologically relevant problems. Since other feasible direction methods are losing popularity, SQP is still a good option.

References

[1] M.S.Bazara, H.D.Sherall and C.M.Shetty, *Nonlinear Programming Theory andAlgorithms* 2^{nd} *edition* New Jersey, (1993).

[2] W.Hock and K.Schittkowski, *A comparative performance evaluation of 27 nonlinear programming codes*, Computing, 30(1983), 335-358. 56(1985), 1-6.

[3] Jorge Nocedal and Stephen J.Wright, *Numerical Optimization*, Springer operation research.

[4] K.Schittkowski, *Nonlinear Programming Codes*, Lecture Notes in economics and Mathematical Systems, 183(1980).

[5] K.Schittkowski, *The non-linear programming method of Wilson, Han and Powell. Part 1: Convergence analysis*, Numerische Mathematik, 38 (1981), 83-114

[6] S.M.Sinha, *Mathematical Programming Theory and Algorithms.*

[7] R. Fletcher, *An ideal penalty function for constrained optimization*, Journal of the Institute of Mathematics and its Applications, 15(1975), 319-342.

[8] R. Fletcher, *Practical Methods of Optimization*, John Wiley, Chichester, (1987).

[9] F. Harary, *On the group of the decomposition of two graphs*, Duke Math. J. 26(1959), 29-34.

[10] R. Fletcher and CM. Reeves, *Function minimization*, Conjug.

[11] S.M.Sinha, *Mathematical Programming Theory and Algorithms (First edition)*, (2006).

[12] Optimization Toolbox Users Guide in Matlab, 2008 by the MathWorks, Inc. Slides designed by Yajun Wang

[13] Engineering Optimization, Concepts and Applications by Fred van Keulen Matthijs Langelaar CLA H21.1

[14] AMSC 607 / CMSC 764 Advanced Numerical Optimization by Dianne P. OLeary

[15] Advanced Mathematical Programming IE417Lecture 22 by Dr. Ted Ralphs.