PRO 5970 Métodos de Otimização Não Linear

Celma de Oliveira Ribeiro

Aula 11 - 2023

Sequential Quadratic Programming

Departamento de Engenharia de Produção
Universidade de São Paulo

Sequential Quadratic Programming - SQP

Consider the problem:

$$
\begin{array}{llr}
 & \text{minimize } f(x) & \\
\text{s.t} & g_i(x) \leq 0 & \text{i} \in \mathcal{I} = \{1, 2, \ldots m\} \\
 & h_i(x) = 0 & \text{i} \in \mathcal{E} = \{1, 2, \ldots l\}
\end{array}
$$

Definition

x^* is KKT point if there are lagrange multipliers vectors λ^* and μ^*, such that $\begin{bmatrix} x^* & \lambda^* & \mu^* \end{bmatrix}^t$ satisfies:

$$
\begin{array}{lll}
\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) & = & 0 \\
g(x^*) & \leq 0 & \\
h(x^*) & = 0 & \\
\mu^* & \geq 0 & \\
\mu_i g_i(x^*) & = 0 & \forall i \in \mathcal{I}
\end{array}
$$

Assumptions:

1. There exist x^*, λ^* and μ^* satisfying KKT conditions
2. The columns of
   G =
   $$\left[\begin{array}{cccccccc} \nabla h_1(x) & \nabla h_2(x) & \dots & \nabla h_l(x) & \nabla g_{i_1}(x) & \dots & \nabla g_{i_p}(x) \end{array} \right]$$
   are linearly independent, $\forall i_j \in \mathcal{I}$

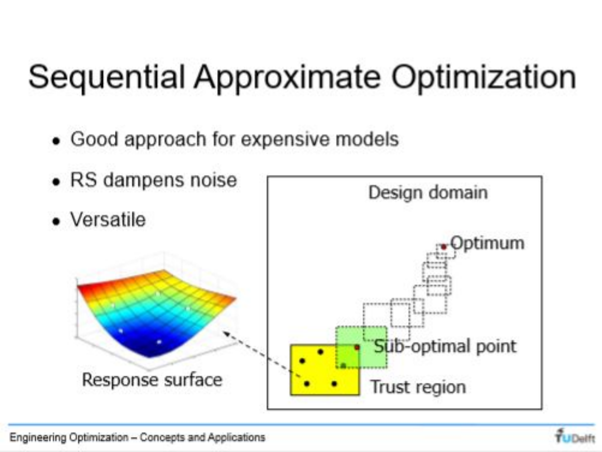

Figure 1: lecture notes from Fred van Keulen and Matthijs Langelaar

Sequential Quadratic Programming (SQP)

- One of the most successful methods for the numerical solution of constrained nonlinear optimization problems.
- It relies on a profound theoretical foundation and provides powerful algorithmic tools for the solution of large-scale technologically relevant problems.

Sequential Quadratic Programming - SQP

Main ideas

- SQP methods are iterative methods that generate a sequence of quadratic optimization problems whose solutions approach the solution of the original problem

Sequential Quadratic Programming - SQP

Main ideas

- SQP methods are iterative methods that generate a sequence of quadratic optimization problems whose solutions approach the solution of the original problem
- At a current iterate x^k, the step to the next iterate is obtained through information generated by solving a quadratic subproblem

Sequential Quadratic Programming - SQP

Main ideas

- SQP methods are iterative methods that generate a sequence of quadratic optimization problems whose solutions approach the solution of the original problem
- At a current iterate x^k, the step to the next iterate is obtained through information generated by solving a quadratic subproblem
- quadratic problems are based on applying KKT conditions to the original problem

Main ideas

- SQP methods are iterative methods that generate a sequence of quadratic optimization problems whose solutions approach the solution of the original problem
- At a current iterate x^k, the step to the next iterate is obtained through information generated by solving a quadratic subproblem
- quadratic problems are based on applying KKT conditions to the original problem
- Minimize a quadratic approximation of the Lagrangian function with respect to linear approximation of the constraints

Idea of algorithms

S1 Choose initial point x_0 and initial lagrange multiplier estimates λ_0

Idea of algorithms

S1 Choose initial point x_0 and initial lagrange multiplier estimates λ_0

S2 Obtain matrices for QP subproblem

Sequential Quadratic Programming - SQP

Idea of algorithms

S1 Choose initial point x_0 and initial lagrange multiplier estimates λ_0

S2 Obtain matrices for QP subproblem

S3 Solve QP subproblem to find a feasible direction d_k

Idea of algorithms

S1 Choose initial point x_0 and initial lagrange multiplier estimates λ_0

S2 Obtain matrices for QP subproblem

S3 Solve QP subproblem to find a feasible direction d_k

S4 $x_{k+1} \leftarrow x_k + \alpha_k d_k$

Idea of algorithms

S1 Choose initial point x_0 and initial lagrange multiplier estimates λ_0

S2 Obtain matrices for QP subproblem

S3 Solve QP subproblem to find a feasible direction d_k

S4 $x_{k+1} \leftarrow x_k + \alpha_k d_k$

S5 If x_{k+1} is not optimal go to S2

The QP subproblems to be solved in each iteration step should reflect the local properties of the NLP with respect to the current iterate x_k.

A natural idea:

- replace objective function by its local quadratic approximation

$$f(x) = f(x_k) + \nabla f(x_k)^t (x - x_k) + \frac{1}{2}(x - x_k)^t \nabla^2 f(x_k)(x - x_k)$$

Sequential Quadratic Programming - SQP

The QP subproblems to be solved in each iteration step should reflect the local properties of the NLP with respect to the current iterate x_k.

A natural idea:

- replace objective function by its local quadratic approximation

$$f(x) = f(x_k) + \nabla f(x_k)^t (x - x_k) + \frac{1}{2}(x - x_k)^t \nabla^2 f(x_k)(x - x_k)$$

- replace constraint functions g and h by their local afine approximations

$$g(x) \approx g(x_k) + \nabla g(x_k)^t (x - x_k)$$

$$h(x) \approx h(x_k) + \nabla h(x_k)^t (x - x_k)$$

Example

Consider the following,

$$\nabla f(x) = \begin{bmatrix} 3 \\ 3 \end{bmatrix},$$

$$\nabla^2 f(x) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$\nabla h_1(x) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \nabla h_2(x) = \begin{bmatrix} 7 \\ 2 \end{bmatrix} \quad \nabla g_1(x) = \begin{bmatrix} 13 \\ 21 \end{bmatrix}$$

Write an approximation of the problem

$$\begin{array}{lll} & \text{minimize } f(x) & \\ \text{s.t} & g_i(x) \leq 0 & \text{i} \in \mathcal{I} = \{1, 2, \ldots m\} \\ & h_i(x) = 0 & \text{i} \in \mathcal{E} = \{1, 2, \ldots l\} \end{array}$$

Assume : $f(x^*) = 100,\ h_1(x^*) = 4,\ h_2(x^*) = 1 - 1,\ g_1(x^*) = 11,$

Sequential Quadratic Programming - SQP

Idea: Consider

$$
\begin{aligned}
&\min_{d\in\mathbb{R}^n} \frac{1}{2} d^t \nabla^2 f(x_k) d + \nabla f(x_k)^t d \\
s.t \quad &\nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\
&\nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E}
\end{aligned}
$$

Note that $d = (x - x^k)$

Instead of analysing the original problem, consider the model related to the Lagrangian $\mathcal{L}$, where λ^* and μ^* are the Lagrangian multipliers associated with the original problem.

$$
\begin{array}{llll}
 & \text{minimize } \mathcal{L}\left(x, \lambda^*, \mu^*\right) & & \\
\text{s.t} & g_i(x) \leq 0 & \text{i} \in \mathcal{I} = \{1, 2, \ldots m\} \\
 & h_i(x) = 0 & \text{i} \in \mathcal{E} = \{1, 2, \ldots l\}
\end{array}
$$

Instead of analysing the original problem, consider the model related to the Lagrangian $\mathcal{L}$, where λ^* and μ^* are the Lagrangian multipliers associated with the original problem.

$$
\begin{array}{llll}
 & \text{minimize } \mathcal{L}\left(x, \lambda^*, \mu^*\right) & & \\
\text{s.t} & g_i(x) \leq 0 & & \text{i} \in \mathcal{I} = \{1, 2, \ldots m\} \\
 & h_i(x) = 0 & & \text{i} \in \mathcal{E} = \{1, 2, \ldots l\}
\end{array}
$$

Although the optimal multipliers are not known, approximations λ^k, μ^k to the multipliers can be maintained as part of the iterative process.

Then given a current iterate, x^k, λ^k, μ^k the quadratic Taylor series approximation in x for the Lagrangian is

$$\mathcal{L}\left(x^k, \lambda^k, \mu^k\right) + \nabla\mathcal{L}(x_k, \mu_k, \lambda_k)^t d + \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d$$

Then given a current iterate, x^k, λ^k, μ^k the quadratic Taylor series approximation in x for the Lagrangian is

$$\mathcal{L}\left(x^k, \lambda^k, \mu^k\right) + \nabla \mathcal{L}(x_k, \mu_k, \lambda_k)^t d + \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d$$

Then we solve:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d + \nabla \mathcal{L}(x_k, \mu_k, \lambda_k)^t d$$

$$\begin{array}{ll} s.t & \nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\ & \nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E} \end{array}$$

$d = (x - x^k)$

Equality constrained problems

Let's consider equality constraints

$$\text{minimize } f(x)$$
$$h_i(x) = 0 \qquad \text{i} \in \mathcal{E} = \{1, 2, \ldots l\}$$

Necessary conditions for a constrained optimal solution are:

$$\nabla f(x) + \sum_{i=1}^{l} \lambda_i h_i(x) = 0$$
$$h_i(x) = 0, i \in \mathcal{E} = \{1, 2, \ldots l\}$$

Equality constrained problems

When we only have equality constraints, we do not have to worry about complementary slackness which makes things simpler...

In this case

$$\min_{x\in\mathbb{R}^n} \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d + \nabla \mathcal{L}(x_k, \mu_k, \lambda_k)^t d$$

$$s.t \quad \nabla h_j(x_k)^t d + h_j(x_k) = 0 \;\; j \in \mathcal{E}$$

is equivalent to

$$\min_{x\in\mathbb{R}^n} \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d + \nabla f(x_k)^t d$$

$$s.t \quad \nabla h_j(x_k)^t d + h_j(x_k) = 0 \;\; j \in \mathcal{E}$$

Explain based on the previous example!

Hint: $\nabla h_j(x_k)^t d = -h_j(x_k) \; j \in \mathcal{E}$ is constant

Equality constrained problems

The SQP Approximation

$$f(x) \approx f(x_k) + \nabla f(x_k)^t d + \frac{1}{2} d^t \nabla_x^2 L(x_k, \lambda_k) d, d = x - x_k$$

$$h(x) \approx h_i(x_k) + \nabla h_i(x_k)^t d = 0 \quad i \in \mathcal{E}$$

Solve the KKT equations for this problem (a linear problem!)

Equality constrained problems

SQP problem at iteration k

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d + \nabla f(x_k)^t d$$

$$s.t \quad \nabla h_j(x_k)^t d + h_j(x_k) = 0 \;\; j \in \mathcal{E}$$

We know how to solve quadratic problems!

Equality constrained problems

SQP problem at iteration k

$$\min_{x\in\mathbb{R}^n} \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d + \nabla f(x_k)^t d$$

$$s.t \quad \nabla h_j(x_k)^t d + h_j(x_k) = 0 \;\; j \in \mathcal{E}$$

We know how to solve quadratic problems!

Use KKT

Equality constrained problems

SQP problem at iteration k

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t \nabla^2 \mathcal{L}(x_k, \mu_k, \lambda_k) d + \nabla f(x_k)^t d$$

$$s.t \quad \nabla h_j(x_k)^t d + h_j(x_k) = 0 \;\; j \in \mathcal{E}$$

We know how to solve quadratic problems!

Use KKT

Solve the following system:

$$\begin{bmatrix} \nabla^2 \mathcal{L}(x_k) & \nabla h((x_k) \\ \nabla^t h((x_k) & 0 \end{bmatrix} \begin{bmatrix} \mathrm{d}_k \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} -\nabla f(x_k) - \nabla h(x_k)\lambda_k \\ -h(x_k) \end{bmatrix}$$

Equality constrained problems

The SQP Approximation

Example: Suppose our approximation is the following,

$$f(x) \approx 3 + \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

$$h(x) \approx 5 + \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

Exercice: Write the lagrangian and the KkT conditions

Equality constrained problems

The KKT equations for this approximation are

$$\begin{bmatrix} 3 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} - \lambda \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 0$$

$$5 + \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = 0$$

The solution is,

$$\begin{bmatrix} d_1 \\ d_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} -2.6 \\ -0.8 \\ 0.4 \end{bmatrix}$$

- This calculation represents the main step in an iteration of the SQP algorithm which solves a sequence of quadratic programs.
- To continue, add d to the current x, update the Lagrangian Hessian, make a new approximation, solve for that solution, and continue iterating in this fashion.

Example

Attention: this approximation considers the gradient of the Lagragian

Consider the following problem:

$$\begin{array}{ll} & \text{minimize } (x_2 - x_1)^4 + (x_1 + x_2 - x_3)^2 \\ \text{s.t.} & \qquad 2x_1 + 3x_2 + 5x_3 = 10 \end{array}$$

Formulate the SQP subproblem for the candidate point $x^k = (2; 2; 0)^t$ and the candidate Lagrange multiplier $\lambda^k = 2$.

The general form of NLO problem is

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & h_j(x) = 0, \ j = 1, \ldots, m. \end{array}$$

The Lagrange function is

$$L(x,y) = f(x) + \sum_{j=1}^{m} y_j h_j(x)$$

where $y_j = \mathbb{R}$, $j = 1, \ldots, m$.

In our case the Lagrange function is

$$L(x,y) = (x_2 - x_1)^4 + (x_1 + x_2 - x_3)^2 + y(2x_1 + 3x_2 + 5x_3 - 10)$$

The gradient and Hessian of the Lagrange function with respect to x are:

$$\nabla_x L(x,y) = \begin{pmatrix} 4(x_1 - x_2)^3 + 2(x_1 + x_2 - x_3) + 2y \\ 4(x_2 - x_1)^3 + 2(x_1 + x_2 - x_3) + 3y \\ -2(x_1 + x_2 - x_3) + 5y \end{pmatrix}$$

$$\nabla^2_{xx} L(x,y) = \begin{pmatrix} 12(x_1 - x_2)^2 + 2 & -12(x_1 - x_2)^2 + 2 & -2 \\ -12(x_1 - x_2)^2 + 2 & 12(x_1 - x_2)^2 + 2 & -2 \\ -2 & -2 & 2 \end{pmatrix}$$

The SQP subproblem is

$$\begin{array}{ll} \min\limits_{\triangle x} & \frac{1}{2}\triangle x^T \nabla^2_{xx} L(x,y)\triangle x + \nabla_x L(x,y)\triangle x \\ \text{s.t.} & \nabla H(x)\triangle x = -H(x). \end{array}$$

where $H(x) = (h_1(x), \ldots, h_m(x))^T$.

Substituting $x = (2,2,0)^T$ and $y = 2$, we get the SQP subproblem

$$\begin{array}{ll} \min\limits_{\triangle x} & \frac{1}{2}\triangle x^T \begin{pmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 2 \end{pmatrix} \triangle x + (\, 12 \;\; 14 \;\; 2\,)\triangle x \\ \text{s.t.} & (\, 2 \;\; 3 \;\; 5\,)\triangle x = 0. \end{array}$$

where $\nabla H(x) = (\, 2 \;\; 3 \;\; 5\,)$ and $H(x) = 0$.

Equality constrained problems

Rudimentary SQP (equality)

Inicialization

$k \leftarrow 1$

Select a starting solution (x_k, λ_k)

Main step

Solve the quadratic problem $QP(x_k, \lambda_k)$ to obtain a solution d_k along with a set of Lagrange multipliers λ_{k+1}

- if $d_k = 0$ stop (x_k, λ_{k+1}) satisfies KKT for the original problem
- Otherwise $x_{k+1} \leftarrow x_k + d_k$

Equality and inequality constrained problems

ref: Bazaraa, ex 10.24

We can consider the inclusion of inequality constraints $g_i(x) \leq 0$, $i \in \{1, 2, ...m\}$ in the problem

The quadratic subproblem $Q(x_k, \mu_k, \lambda_k)$ is given as

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t \nabla_k^2 \mathcal{L}(x_k, \lambda_k, \mu_k) d + \nabla f(x_k)^t d$$

$$\begin{array}{ll} s.t & \nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\ & \nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E} \end{array}$$

Equality and inequality constrained problems

ref: Bazaraa

The KKT conditions for this problem require that, in addition to feasibility, the complementary conditions to be satisfied for Lagrange multipliers $(\mu_{k+1}, \lambda_{k+1})$

$$\mu_i[\nabla g_i(x_k)^t d + g_i(x_k)] = 0 \quad i \in \mathcal{I}$$

$$\mu \geq 0$$

Hence, if d_k solves $Q(x_k, \mu_k, \lambda_k)$ with lagrange multipliers $(\mu_{k+1}, \lambda_{k+1})$

- if $d_k = 0$, then x_k, along with $(\mu_{k+1}, \lambda_{k+1})$ yields a KKT solution for the original problem
- Otherwise set $x_{k+1} \leftarrow x + k + d_k$, increment k by 1, and repeat the process.

It can be shown that under second-order sufficiency conditions, if (x_k, μ_k, λ_k) is initialized sufficiently close to the optimal solution, the iterative process will converge quadratically to the solution.

See example 10.4.3 from Bazaraa!

Equality and inequality constrained problems

Quasi-Newton approximations

Disadvantage of the SQP method: it requires second order derivatives to be calculated and the Hessian might not be positive definite.

Equality and inequality constrained problems

Quasi-Newton approximations

Disadvantage of the SQP method: it requires second order derivatives to be calculated and the Hessian might not be positive definite.

- We can adopt an approach similar to the unconstrained case, considering approximations H_k of the hessian $\nabla_x^2 \mathcal{L}(x_k, \lambda_k, \mu_k)$
- Create a quadratic approximation to the Lagrangian and linear approximations to the constraints

Equality and inequality constrained problems

Quasi-Newton approximations

Disadvantage of the SQP method: it requires second order derivatives to be calculated and the Hessian might not be positive definite.

- We can adopt an approach similar to the unconstrained case, considering approximations H_k of the hessian $\nabla_x^2 \mathcal{L}(x_k, \lambda_k, \mu_k)$
- Create a quadratic approximation to the Lagrangian and linear approximations to the constraints
- At the kth iteration solve a quadratic subproblem (QP) of the form

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t H_k d + \nabla f(x_k)^t d$$

$$\begin{array}{ll} s.t & \nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\ & \nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E} \end{array}$$

  where d is the search direction and H_k is a positive definite approximation to the Hessian matrix of Lagrangian function of the original problem

Equality and inequality constrained problems

Quasi-Newton approximations

Disadvantage of the SQP method: it requires second order derivatives to be calculated and the Hessian might not be positive definite.

- We can adopt an approach similar to the unconstrained case, considering approximations H_k of the hessian $\nabla_x^2 \mathcal{L}(x_k, \lambda_k, \mu_k)$
- Create a quadratic approximation to the Lagrangian and linear approximations to the constraints
- At the kth iteration solve a quadratic subproblem (QP) of the form

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t H_k d + \nabla f(x_k)^t d$$

$$\begin{array}{ll} s.t & \nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\ & \nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E} \end{array}$$

  where d is the search direction and H_k is a positive definite approximation to the Hessian matrix of Lagrangian function of the original problem
- Solve the quadratic problem to find the search direction, d_k, used to generate a new iterate $x_{k+1} = x_k + \alpha_k d_k$

Equality and inequality constrained problems

Quasi-Newton approximations

Disadvantage of the SQP method: it requires second order derivatives to be calculated and the Hessian might not be positive definite.

- We can adopt an approach similar to the unconstrained case, considering approximations H_k of the hessian $\nabla_x^2 \mathcal{L}(x_k, \lambda_k, \mu_k)$
- Create a quadratic approximation to the Lagrangian and linear approximations to the constraints
- At the kth iteration solve a quadratic subproblem (QP) of the form

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t H_k d + \nabla f(x_k)^t d$$

$$\begin{array}{ll} s.t & \nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\ & \nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E} \end{array}$$

  where d is the search direction and H_k is a positive definite approximation to the Hessian matrix of Lagrangian function of the original problem
- Solve the quadratic problem to find the search direction, d_k, used to generate a new iterate $x_{k+1} = x_k + \alpha_k d_k$
- Perform line search to find α_k

Equality and inequality constrained problems

Quasi-Newton approximations

Disadvantage of the SQP method: it requires second order derivatives to be calculated and the Hessian might not be positive definite.

- We can adopt an approach similar to the unconstrained case, considering approximations H_k of the hessian $\nabla_x^2 \mathcal{L}(x_k, \lambda_k, \mu_k)$
- Create a quadratic approximation to the Lagrangian and linear approximations to the constraints
- At the kth iteration solve a quadratic subproblem (QP) of the form

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} d^t H_k d + \nabla f(x_k)^t d$$

$$\begin{array}{ll} s.t & \nabla g_i(x_k)^t d + g_i(x_k) \leq 0 \quad i \in \mathcal{I} \\ & \nabla h_j(x_k)^t d + h_j(x_k) = 0 \quad j \in \mathcal{E} \end{array}$$

  where d is the search direction and H_k is a positive definite approximation to the Hessian matrix of Lagrangian function of the original problem
- Solve the quadratic problem to find the search direction, d_k, used to generate a new iterate $x_{k+1} = x_k + \alpha_k d_k$
- Perform line search to find α_k
- If optimality not achieved update the approximation to the Lagrangian

Some important issues

- The solution d_k of the quadratic program (QP) can be used to generate a new iterate x^{k+1} by taking a step from x^k in the direction of d_k.

Some important issues

- The solution d_k of the quadratic program (QP) can be used to generate a new iterate x^{k+1} by taking a step from x^k in the direction of d_k.
- To continue to the next iteration, new estimates for the Lagrange multipliers are needed.

Some important issues

- The solution d_k of the quadratic program (QP) can be used to generate a new iterate x^{k+1} by taking a step from x^k in the direction of d_k.
- To continue to the next iteration, new estimates for the Lagrange multipliers are needed.
- One obvious approach: use the optimal multipliers of the quadratic subproblem
  Let the optimal multipliers of QP be λ_{qp} and μ_{qp} and setting $d_\lambda = \lambda_{qp} - \lambda^k$ and $d_\mu = \mu_{qp} - \mu^k$, the updates will be :
  - $x^{k+1} \leftarrow x^k + \alpha d$
  - $\lambda^{k+1} \leftarrow \lambda^k + \alpha d_\lambda$
  - $\mu^{k+1} \leftarrow \mu^k + \alpha d_\mu$

  for some selection of the steplength parameter α

SQP gradually enforces feasibility of the constraints as part of the KT equations

Advantages and disadvantages

- One of the most successful methods for solving NLP
- It is not an algorithm but rather a conceptual method
- Usually violates non linear constraints until convergence, often by large amounts
- Requires a good QP solver

Linear constraints

Not presented in 2023

Consider the problem (PLC)

$$
\begin{aligned}
& \text{minimize } f(x) \\
\text{s.t} \quad & Ax \leq b \\
& Ex = e
\end{aligned}
$$

Lemma

Let x be a feasible solution for the PLC problem. Suppose $A_1 x = b_1$ and $A_2 x < b_2$. Decompose matrix A and vector b as:

$A^t = \begin{bmatrix} A_1^t & A_2^t \end{bmatrix}$ and $b^t = \begin{bmatrix} b_1^t & b_2^t \end{bmatrix}$

Then a non zero vector d is a feasible direction at $x \Leftrightarrow A_1 d \leq 0$ and $Ed = 0$.

If $\nabla f(x)^t d < 0$, then d is a descent direction.

Example

Consider

$$\begin{array}{ll} \text{minimize} & (x_1-2)^2+(x_2-6)^2 \\ \text{s.t} & -x_1+2x_2 \leq 4 \\ & 3x_1+2x_2 \leq 12 \\ & -x_1 \leq 0 \\ & -x_2 \leq 0 \end{array}$$

Let $x = \begin{bmatrix} 2 & 3 \end{bmatrix}^t$

Find the set of descent directions at x

$$A_1 = \begin{bmatrix} -1 & 2 \\ 3 & 2 \end{bmatrix}$$

Vector d is a feasible solution $\Leftrightarrow \begin{cases} -1d_1 + 2d_2 \leq 0 \\ 3d_1 + 2d_2 \leq 0 \end{cases}$

Descent directions $\{d \in \mathbb{R}^2 | -8d_1 + 2d_2 < 0\}$

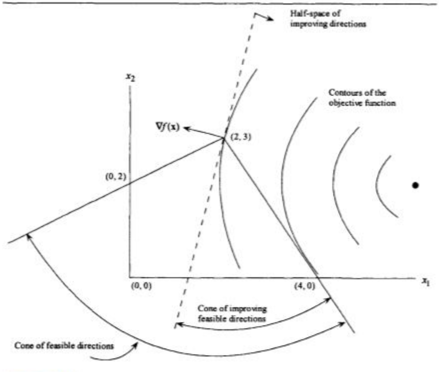

Figure 10.1 Improving feasible directions.

Generating Descent Feasible directions

Solve the problem:

$$\begin{array}{ll} \text{minimize} & \nabla f(x)^t d \\ \text{s.t} & A_1 d \leq 0 \\ & Ed = 0 \\ & -1 \leq d \leq \mathbf{1} \end{array}$$

x is a KKT point $\Leftrightarrow$ the optimal solution of the problem is equal to zero

Line search

Given x_k, the next point is $x_{k+1} \leftarrow x_k + \lambda_k d_k$. The value λ_k is obtained through:

$$\begin{array}{llll} \text{minimize} & f(x_k + \lambda_k d_k) & & \\ \text{s.t} & A(x_k + \lambda_k d_k) & \leq b & \\ & E(x_k + \lambda_k d_k) & = e & \textbf{Redundant!} \\ & \lambda_k & \geq 0 & \end{array}$$

Note that

$$A_1 x_k + \lambda_k A_1 d_k = b_1 + \underbrace{\lambda_k}_{\geq 0} \underbrace{A_1 d_k}_{\leq 0} \leq b_1 \ \textbf{Redundant!}$$

$$\lambda_k A_2 d_k = \underbrace{b_2 - A_2 x_k}_{\geq 0}$$

Line search

The line search then reduces to:

$$\begin{array}{cc} \text{minimize} & f(x_k + \lambda d_k) \\ \text{s.t} & 0 \leq \lambda \leq \lambda_{max} \end{array}$$

with $\lambda_{max} = \begin{cases} min\left\{\frac{\hat{p}_i}{\hat{q}_i} | \hat{q}_i > 0\right\} & \hat{q} \not\leq 0 \\ \infty & \hat{q} \leq 0 \end{cases}$

$$\hat{p} = b_2 - A_2 x_k$$

$$\hat{q} = A_2 d_k$$

ZOUTENDIJK algorithm (Linear constraints)

Step 0

- Find a starting feasible solution x_1 with $Ax_1 \leq b$ and $Ex_1 = e$
- $k \leftarrow 1$

Step 1

- Given x_k solve
  minimize$\left\{ \nabla f(x_k)^t d | A_1 d \leq 0 \quad Ed = 0 \quad -1 \leq d \leq \mathbf{1} \right\}$
- if $\nabla f(x_k)^t d = 0$ stop x_k is KKT point.
  Otherwise, go to step 2

Step 2

- Let λ_k be an optimal solution to the line search problem: min $_{0 \leq \lambda \leq \lambda_{max}} \{ f(x_k + \lambda d_k) \}$
- $x_{k+1} \leftarrow x_k + \lambda_k d_k$
- Identify the new set of binding constraints at x_{k+1}, and update A_1 and A_2 accordingly.
  $k \leftarrow k + 1$ Go to step 1.

Zoutendijk's method

- Subproblem linear: efficiently solved
- Determine active set before solving subproblem!
- When $\nabla f(x_k)^t d = 0$ KKT point found
- Method needs a feasible starting point.
- Convergence is not generally guaranteed (modification of the method assures convergence)

Example 1

Consider

$$\begin{array}{ll} \text{minimize} & 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2 \\ \text{s.t} & x_1 + x_2 \leq 2 \\ & x_1 + 5x_2 \leq 5 \\ & -x_1 \leq 0 \\ & -x_2 \leq 0 \end{array}$$

Solve the problem considering the initial point $x^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$