A FEASIBLE POINT ALGORITHM FOR NONLINEAR CONSTRAINED OPTIMIZATION with Applications to Parameter Identification in Structural Mechanics

José Herskovits[1]

[1] COPPE - Federal University of Rio de Janeiro, Mechanical Engineering Department
Rio de Janeiro, Brazil.

IMPA, Rio de Janeiro, Brazil
October 2017

Introduction

We consider the nonlinear constrained optimization program:

$$\begin{cases} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & g(x) \leqslant 0, g \in \mathbb{R}^m \\ \text{and} & h(x) = 0, h \in \mathbb{R}^p \end{cases}$$

where:

$f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$

$f(x)$, $g(x)$ and $h(x)$ are smooth functions, not necessarily convex.

Introduction

- We present a general approach for interior point algorithms to solve the nonlinear constrained optimization problem.
- Given an initial estimate x of at the interior of the inequality constraints, these algorithms define a sequence of interior points with the objective reduced at each iteration.
- By this technique, first order, quasi Newton or Newton algorithms can be obtained.

Introduction

In this talk, we present:

- A FEASIBLE DIRECTION ALGORITHM At each point a descent feasible direction is obtained. Then, an inaccurate line search is done to get a new interior point with a reasonable decrease of the objective.
- A FEASIBLE ARC ALGORITHM The line search is done along an arc.

Introduction

The present approach is:

- Simple to code, strong and efficient.
- It does not involve penalty functions, active set strategies or Quadratic Programming subproblems.
- It merely requires to solve two linear systems with the same matrix at each iteration and to perform an inaccurate line search.
- In practical applications, more efficient algorithms can be obtained by taking advantage of the structure of the problem and particularities of the functions in it.

Nonlinear constrained optimization

- Our approach is based on FDIPA - The Interior Point Algorithm for Standart Nonlinear Constrained Optimization.
- Following we shall describe FDIPA and the basic ideas involved in it.
- Herskovits, J., “A Feasible Directions Interior Point Technique For Nonlinear Optimization”, Journal of Optimization Theory and Algorithms, 1998.
- Herskovits, J., “A two-stage feasible directions algorithm for nonlinear constrained optimization”, Mathematical Programming, 1986.

About FDIPA

- FDIPA is a general technique to solve nonlinear constrained optimization problems.
- Requires an initial point at the interior of the inequality constraints and generates a sequence of interior points.
- When the problem has only inequality constraints the objective function is reduced at each iteration.
- FDIPA only requires the solution of 2 linear systems with the same matrix at each iteration.
- FDIPA is very robust and it no requires parameters tuning.

About FDIPA

We describe now FDIPA and discuss the ideas involved in it, in the framework of the inequality constrained problem:

$$\begin{cases} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & g(x) \leqslant 0 \end{cases}$$

Let be the feasible set:

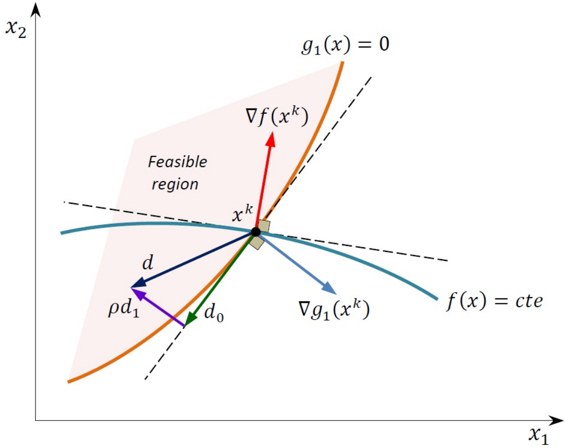
$$\Omega \equiv \left\{ x \in \mathbb{R}^n : g(x) \leqslant 0 \right\}$$

Definitions

- $d \in \mathbb{R}^n$ at $x \in \mathbb{R}^n$ is a descent direction for a smooth function $\phi(x) : \mathbb{R}^n \to \mathbb{R}$ if $d^T \nabla \phi(x) < 0$.
- $d \in \mathbb{R}^n$ is a feasible direction for the problem, at $x \in \Omega$, if for some $\theta(x) > 0$ we have $x + td \in \Omega$ for all $t \in [0, \theta(x)]$.
- A vector field $d(x)$ defined on Ω is said to be a uniformly feasible directions field of the problem (2), if there exists a step length $\tau > 0$ such that $x + td(x) \in \Omega$ for all $t \in [0, \tau]$ and for all $x \in \Omega$. That is, $\tau \leq \theta(x)$ for $x \in \Omega$.

About FDIPA

- Search Direction



FDIPA – Feasible Direction Interior Point Algorithm

- **Parameter.** $\xi \in (0,1)$, $\eta \in (0,1)$, $\varphi > 0$ e $\nu \in (0,1)$.
- **Data.** $x \in int(\Omega_a)$, $\lambda \in \mathbb{R}^m_+$ and $B \in \mathbb{R}^{n\times n}$, where the initial quasi-Newton matrix is Symmetric and Positive Definite.
- **Step 1.** Computation of the search direction d.

  (i) Solve the following linear systems to obtain $d_0, d_1 \in \mathbb{R}^n$ and $\lambda_0, \lambda_1 \in \mathbb{R}^m$.

$$\begin{bmatrix} B & \nabla g(x) \\ \Lambda \nabla g^\top(x) & G \end{bmatrix} \begin{bmatrix} d_0 & d_1 \\ \lambda_0 & \lambda_1 \end{bmatrix} = \begin{bmatrix} -\nabla f(x) & 0 \\ 0 & -\lambda \end{bmatrix}$$

  If $d_0 = 0$, stop.

  (ii) If $d_1^\top \nabla f(x) > 0$, compute $\rho = \min \left\{ \varphi \|d_0\|^2, (\xi - 1)\dfrac{d_0^\top \nabla f(x)}{d_1^\top \nabla f(x)} \right\}$

  Otherwise, compute $\rho = \varphi \|d_0\|^2$

  (iii) Compute the search direction $d = d_0 + \rho d_1$.

FDIPA – Feasible Direction Interior Point Algorithm

- **Step 2.** Line Search.

  Find t, the first element of $\{1, \nu, \nu^2, \nu^3 \dots\}$ such that

  $$f(x + td) \leqslant f(x) + t.\eta.d^\top \nabla f(x)$$

  and

  $$\begin{aligned} g_i(x + td) < 0, &\quad \bar{\lambda}_i \geq 0, \\ g_i(x + td) \leq g_i(x), &\quad \bar{\lambda}_i < 0. \end{aligned}$$

- **Step 3.** Updates.

  (i) Update the new point by $x = x + td$.

  (ii) Define a new value for $B \in \mathbb{R}^{m \times m}$ symmetric and positive definite.

  (iii) Define a new value for $\lambda \in \mathbb{R}^m_+$

  (iv) Go to Step 1.

About FDIPA

- At each point FDIPA computes first a **Search Direction** that is a **Feasible Descent Direction** of the problem
- Then, through a line search procedure, **a new feasible point with a lower cost**, is obtained
- In fact, **the search directions constitute a Uniformly Feasible Directions Field**

About FDIPA

Karush-Kuhn-Tucker optimality conditions:

If x is a local minimum, then

$$\begin{aligned} \nabla f(x) + \nabla g(x)\lambda &= 0 \\ G(x)\lambda &= 0 \\ g(x) &\leqslant 0 \\ \lambda &\geqslant 0 \end{aligned}$$

where:

$\lambda \in \mathbb{R}^m$ are the dual variables, and

$G(x)$ diagonal matrix with $G_{ii}(x) = g_i(x)$.

In the present approach, we look for (x, λ) that satisfy KKT conditions.

Assumptions

- There exists a real number a such that the set $\Omega_a \equiv \{x \in \Omega : f(x) \leqslant a\}$ is compact and has an interior Ω_a^0.
- Each $x \in \Omega_a^0$ satisfy $g(x) < 0$.
- The functions f and g are continuously differentiable in Ω_a and their derivatives satisfy a Lipschitz condition.
- (Regularity Condition). For all Stationary Point $x^* \in \Omega_a$, the vectors $\nabla g_i(x^*)$, for i such that $g_i(x^*) = 0$, are linearly independent.

About FDIPA

We propose Newton like iterations to solve the equations in KKT conditions:

$$\begin{aligned} \nabla f(x) + \nabla g(x)\lambda &= 0 \\ G(x)\lambda &= 0 \end{aligned}$$

in such a way that each iterate satisfies the inequations:

$$\begin{aligned} g(x) &\leqslant 0 \\ \lambda &\geqslant 0 \end{aligned}$$

We define a function ψ such that:

$$\psi(x, \lambda) = 0 \iff \begin{cases} \nabla_x L(x, \lambda) = 0 \\ \Lambda G(x) = 0 \end{cases}$$

The Jacobian matrix of ψ is:

$$\begin{bmatrix} B & \nabla g(x) \\ \Lambda \nabla g^T(x) & G(x) \end{bmatrix}$$

About FDIPA

A Newton-like Iteration in (x, λ) for the equations in KKT condition is:

$$\begin{bmatrix} B & \nabla g(x) \\ \Lambda \nabla g^T(x) & G(x) \end{bmatrix} \begin{bmatrix} x_0 - x \\ \lambda_0 - \lambda \end{bmatrix} = - \begin{bmatrix} \nabla f(x) + \nabla g(x)\lambda \\ G(x)\lambda \end{bmatrix}$$

where:

(x, λ) is the present point,

(x_0, λ_0) is a new estimate.

Λ is a diagonal matrix such that $\Lambda_{ii} = \lambda_i$.

About FDIPA

We can take:

- $B = \nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i \nabla^2 g(x)$: a Newton's Method;
- B : a quase-Newton approx.: Quasi-Newton;
- $B = I$: a First order method.

We define now the vector d_0 in the primal space, as

$$d_0 = x_0 - x$$

Then, we have:

$$\begin{aligned} Bd_0 + \nabla g(x)\lambda_0 &= -\nabla f(x) \\ \Lambda \nabla g^T(x) d_0 + G(x)\lambda_0 &= 0 \end{aligned}$$

About FDIPA

We prove that, if;

- B is Positive Definite;
- $\lambda > 0$.

and

- $g(x) \leqslant 0$;

then:

- The system has an unique solution;
- d_0 is a descent direction of $f(x)$.

About FDIPA

However, d_0 is not always a feasible direction.

In fact,

$$\Lambda \nabla g^T(x) d_0 + G(x)\lambda_0 = 0$$

is equivalent to:

$$\lambda_i \nabla g_i^T(x) d_0 + g_i(x)\lambda_{0i}; \quad i = 1, \ldots, m$$

Thus, d_0 is not always feasible since is tangent to the active constraints.

About FDIPA

Then, to obtain a feasible direction, a negative number is added in the right side:

$$\lambda_i \nabla g_i^T(x) d + g_i(x) \overline{\lambda}_i = -\rho \lambda_i \omega_i, \quad i = 1, \ldots, m,$$

and we get a new perturbed system:

$$\begin{aligned} Bd + \nabla g(x) \overline{\lambda} &= -\nabla f(x) \\ \Lambda \nabla g^T(x) d + G(x) \overline{\lambda} &= -\rho \lambda \end{aligned}$$

where $\rho > 0$.

The negative number in the right hand side produces the effect of bending d_0 to the interior of the feasible region, being the deflection relative to each constraint proportional to ρ.

About FDIPA

- As the deflection is proportional to ρ and d_0 is descent, by establishing upper bounds on ρ, it is possible to ensure that d is a descent direction also.

- Since $d_0^T \nabla f(x) < 0$, we can obtain these bounds by imposing:

$$d^T \nabla f(x) \leqslant \alpha d_0^T \nabla f(x),$$

which implies $d^T \nabla f(x) < 0$.

About FDIPA

Let us consider

$$Bd_0 + \nabla g(x)\lambda_0 = -\nabla f(x)$$
$$\Lambda \nabla g^T(x) d_0 + G(x)\lambda_0 = 0$$

and the auxiliary system of linear equations

$$Bd_1 + \nabla g(x)\lambda_1 = 0$$
$$\Lambda \nabla g^T(x) d_1 + G(x)\lambda_1 = -\lambda$$

About FDIPA

- We have that the roots of

$$B d + \nabla g(x)\overline{\lambda} = -\nabla f(x)$$
$$\Lambda \nabla g^T(x) d + G(x)\overline{\lambda} = -\rho\lambda$$

  are

$$d = d_0 + \rho d_1$$

  and

$$\overline{\lambda} = \lambda_0 + \rho\lambda_1$$

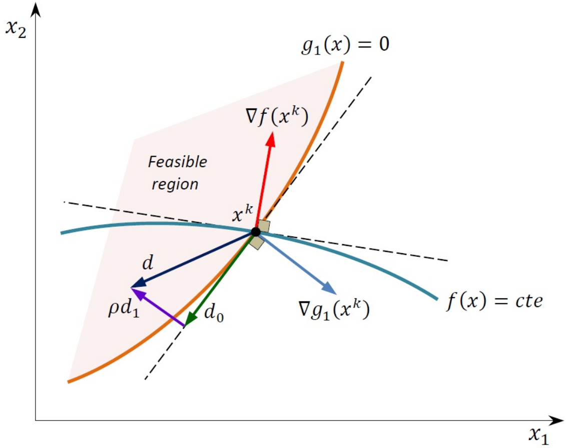
- By substitution of $d = d_0 + \rho d_1$ in $d^T \nabla f(x) \leqslant \alpha d_0^T \nabla f(x)$, we get

$$\rho \leqslant (\alpha - 1)\frac{d_0^T \nabla f(x)}{d^T \nabla f(x)}$$

  - in the case when $d_1^T \nabla f(x) > 0$
  - Otherwise, any $\rho > 0$ holds.

About FDIPA

- Search Direction



About FDIPA

- In fact, the search direction d constitutes an uniformly feasible directions field.
- To find a new primal point, an inaccurate line search is done in the direction of d.
- We look for a new interior point with a satisfactory decrease of the objective.
- Different updating rules can be employed to define a new λ positive.

About the Line Search

- We extend Armijo's Line Search to this kind of interior point algorithms:

  Compute t, the first number of the sequence $\{1, \nu, \nu^2, \nu^3, \ldots\}$, such that:

  $$f(x + td) < f(x) + t\eta \nabla f^T(x)d$$

  and

  $$g_i(x + td) < 0, \, if \ \ \overline{\lambda}_i \geqslant 0$$

  or

  $$g_i(x + td) \leqslant g_i(x), \text{otherwise}$$

- These conditions:
  (i) Ensure a reasonable decrease of the function.
  (ii) Only accept interior points.
  (iii) Avoids saturation of non active constraints.

- In practice we employ more efficient line search procedures:
  (i) An extension of Wolfe's rule, combined with interpolations of the functions.
  (ii) An extension of Goldstein's rule, combined with interpolations of the functions.

Theoretical Results

- First we prove that the algorithm never fails, in particular that the linear systems have an unique solution.
- We prove that any sequence given by the algorithm converges to a KKT point for any way of updating, and, that verify the assumptions above.
- Moreover, converges to Karush-Kuhn- Tucker pair. Depending on the way of updating , global convergence in the dual space can also be obtained.

Theoretical Results

- Several practical applications and test problems were solved very efficiently with this algorithm.
- However for some problems with highly nonlinear constraints the unitary step length is not obtained and the rate of convergence is worst than superlinear.
- This effect is similar to the Maratos' effect for the feasible direction methods and occurs when the feasible direction supports a too short feasible segment.

The Feasible Arc technique avoids this effect.

Theoretical Results

- The arc search technique was first presented by Mayne and Polak, 1976.
- Painer and Tits, 1987, employed it in a feasible SQP algorithm.
- Panier, Tits and Herskovits, 1988 modified FD-IPA by including an arc search.
- All this methods require an additional SQP subproblem.
- With the present approach we only need to solve an additional linear system.

About FAIPA

To include 2nd order information about the constraints:
We compute the search direction d of the Feasible Directions Algorithm, and

$$\tilde{\omega}_i = g_i(x + d) - g_i(x) - \nabla g_i(x) d.$$

In effect:

$$\tilde{\omega}_i \approx \frac{1}{2} d^\top \nabla^2 g_i(x) d$$

That is, $\tilde{\omega}$ is an approximation of the 2nd derivative of $g_i(x)$ along d.

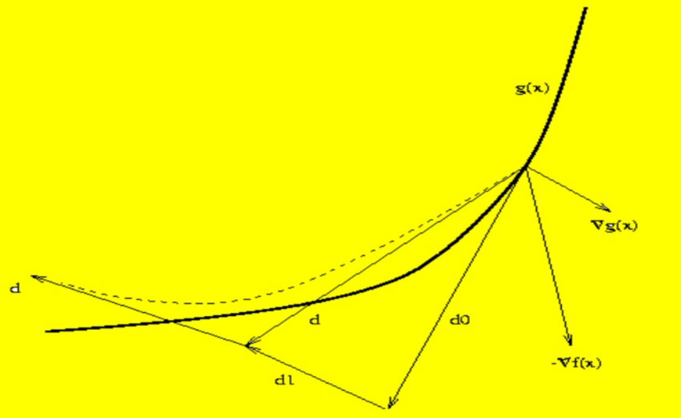
Finally,

i) The 2nd order correction $\tilde{d}$ is computed by solving:

$$\begin{aligned} B\tilde{d} + \nabla g(x)\tilde{\lambda} &= 0; \\ \lambda \nabla g^{\top}(x)\tilde{d} + g(x)\tilde{\lambda} &= -\lambda\tilde{\omega}, \end{aligned}$$

ii) The feasible arc is defined as: $x^{k+1} = x^k + td + t^2\tilde{d}$.

- Search Arc



About FAIPA

We prove that:

i) For any $x \in \Omega$, there is a $\theta(x) > 0$ such that $x = td + t^2\tilde{d}$ for all $t \in [0, \theta(x)]$.

ii) $\theta(x) > 1$ for x near enough of a Karush - Kuhn – Tucker Point.

FAIPA – Feasible Arc Interior Point Algorithm

- **Parameter.** $\xi \in (0,1)$, $\eta \in (0,1)$, $\varphi > 0$ e $\nu \in (0,1)$.
- **Data.** $x \in int(\Omega_a)$, $\omega \in \mathbb{R}^m_+$, $\lambda \in \mathbb{R}^m_+$ and $B \in \mathbb{R}^{n\times n}$, where the initial quasi-Newton matrix is Symmetric and Positive Definite.
- **Step 1.** Computation of the search direction d.

(i) Solve the following linear systems to obtain $d_0, d_1 \in \mathbb{R}^n$ and $\lambda_0, \lambda_1 \in \mathbb{R}^m$.

$$\begin{bmatrix} B & \nabla g(x) \\ \Lambda \nabla g^\top(x) & G \end{bmatrix} \begin{bmatrix} d_0 & d_1 \\ \lambda_0 & \lambda_1 \end{bmatrix} = \begin{bmatrix} -\nabla f(x) & 0 \\ 0 & -\Lambda\omega \end{bmatrix}$$

If $d_0 = 0$, stop.

(ii) If $d_1^\top \nabla f(x) > 0$, compute $\rho = \min\left\{ \varphi \|d_0\|^2, (\xi - 1)\dfrac{d_0^\top \nabla f(x)}{d_1^\top \nabla f(x)} \right\}$

Otherwise, compute $\rho = \varphi \|d_0\|^2$

(iii) Compute the search direction $d = d_0 + \rho d_1$.

- **Step 2.** Compute $\tilde{\omega}^i$ and $\tilde{d}$

  (i) Compute
  $$\tilde{\omega}^i = g_i(x+d) - g_i(x) - \nabla g_i^\top(x)d$$
  (ii) Solve the linear system:
  $$\begin{aligned} B\tilde{d} + \nabla g(x)\tilde{\lambda} &= 0 \\ \Lambda \nabla g^\top(x)\tilde{d} + G(x)\tilde{\lambda} &= \Lambda\tilde{\omega} \end{aligned}$$

FAIPA – Feasible Arc Interior Point Algorithm

- **Step 3.** Arc Search.

  Find t, the first element of $\{1, \nu, \nu^2, \nu^3 \ldots\}$ such that

  $$f(x + td + t^2\tilde{d}) \leqslant f(x) + t.\eta.d^\top \nabla f(x)$$

  and

  $$\begin{aligned} g_i(x + td + t^2\tilde{d}) < 0, &\quad \bar{\lambda}_i \geq 0, \\ g_i(x + td + t^2\tilde{d}) \leq g_i(x), &\quad \bar{\lambda}_i < 0. \end{aligned}$$

- **Step 4.** Updates.

  (i) Update the new point by $x = x + td + t^2\tilde{d}$.

  (ii) Define a new value for $B \in \mathbb{R}^{m\times m}$ symmetric and positive definite.

  (iii) Define a new value for $\lambda \in \mathbb{R}^m_+$

  (iv) Go to Step 1.

Ackowledgements

The authors which to acknowledge to CNPq, FAPERJ and CAPES for the support provided to this research.

Thanks!!!!!