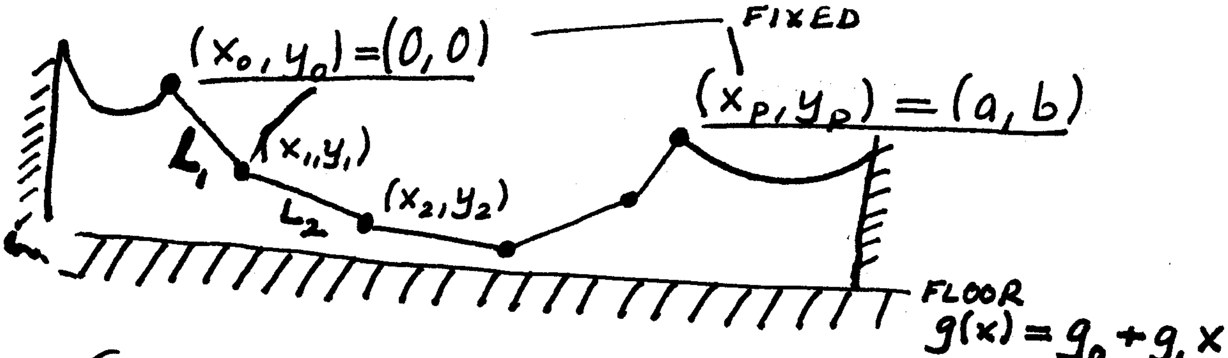
LECTURE 26

NONLINEAR PROGRAMS AND THEIR NUMERICAL SOLUTION

$$
\text{(NP)} \quad \begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & \\ c_i(x) \geq 0 & i = 1, \ldots, m \\ \tilde{c}_j(x) = 0 & j = 1, \ldots, l \end{array}
$$

Hanging Chain Problem



Given

* coordinates of the end-joints $(x_0, y_0) = (0,0)$ and $(x_p, y_p) = (a,b)$
* length of the jth rod $L_j \quad j = 1, \ldots, p$
* all joints must be above the floor, i.e., $g_0 + g_1 x_j \leq y_j$

PROBLEM

(1) p rods are connected at joint points with coordinates (x_j, y_j)
$j = 1, \ldots, p$

(2) The length of jth rod is fixed and equal to L_j. But the coordinates of the jth joint (x_j, y_j) are variables.

(3) Find the joint coordinates minimizing the total potential

Objective Function

$$\phi(x) = \sum_{j=1}^{p} \underbrace{L_j \frac{(y_j + y_{j-1})}{2}}_{\text{Potential of the } j\text{th rod}}$$

VECTOR OF UNKNOWNS

$$X = [x_1 \;\; x_2 \;\; \cdots \;\; x_{p-1} \;\; y_1 \;\; y_2 \;\; \cdots \;\; y_{p-1}]^T \in \mathbb{R}^{2p-2}$$

Equality Constraints

$$j = 1, \ldots, p \qquad L_j^2 = (x_j - x_{j-1})^2 + (y_j - y_{j-1})^2$$

that is the length of the jth rod is L_j.

Equivalently

$$\underbrace{(x_j - x_{j-1})^2 + (y_j - y_{j-1})^2 - L_j^2}_{\tilde{c}_j(x)} = 0$$

Inequality Constraints

$j = 0, \dots, p \qquad g_0 + g_1 x_j \leqslant y_j$

that is the jth joint must be above the floor.

~~Proble~~

Equivalently

$$\underbrace{y_i - g_0 - g_1 x_i}_{c_i(x)} \geqslant 0$$

Problem Formulation

$$\underset{x \in \mathbb{R}^{2p-2}}{\text{minimize}} \quad \sum_{j=1}^{p} L_j \left(\frac{y_j + y_{j-1}}{2} \right)$$

subject to

$$y_i - g_0 - g_1 x_i \geqslant 0 \qquad i = 0, \dots, p$$

$$(x_j - x_{j-1})^2 + (y_j - y_{j-1})^2 - L_j^2 = 0 \quad j = 1, \dots, p$$

(Above in the objective and constraints $x_0 = 0,\ x_p = a,\ y_0 = 0,\ y_p = b$)

Penalty Function Method

First focus on equality constraints only.

$$\text{(NEP)} \quad \begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \\ \text{subject to} \\ \quad \tilde{c}_j(x) = 0 \qquad j = 1, \dots, l \end{array}$$

Basic Idea

* Replace the problem (NEP) by an unconstrained optimization problem
* Penalize the violation of equality constraints

Penalty Function

$$P(x, \mu) = f(x) + \frac{1}{2\mu} \sum_{j=1}^{l} \tilde{c}_j^2(x)$$

where $\mu > 0$ is the penalty parameter.

EXAMPLE

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad x_1 + x_2$$

subject to

$$x_1^2 + x_2^2 = 1$$

Penalty Function

$$P(x, \mu) = x_1 + x_2 + \frac{1}{2\mu}\left(x_1^2 + x_2^2 - 1\right)^2$$

Algorithm

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad P(x, \mu)$$

Solve the unconstrained optimization problem above for various μ

* starting with large μ
* eventually letting $\mu \to 0$

EXAMPLE

$$\text{minimize} \quad x_2$$

$$x_1^2 - x_2 = 1$$

Optimal objective function value

$$(x_*)_2 = (x_*)_1^2 - 1 = -1$$

Penalty function

$$P(x,\mu) = x_2 + \frac{1}{2\mu}\left(x_1^2 - x_2 - 1\right)^2$$

Finding a local min as a function of μ

* $$\nabla_x P(x,\mu) = \begin{bmatrix} 2x_1\left(x_1^2 - x_2 - 1\right)/\mu \\ 1 - \left(x_1^2 - x_2 - 1\right)/\mu \end{bmatrix}$$

* only stationary point

$$\left(x_1(\mu),\ x_2(\mu)\right) = (0, -1-\mu)$$

* Hessian

$$\nabla^2_{xx} P(x,\mu) = \begin{bmatrix} 0 & 0 \\ 0 & 1/\mu \end{bmatrix} \geq 0$$

* Consequently

$$\left(x_1(\mu),\ x_2(\mu)\right) = (0, -1-\mu)$$

is the unique local minimizer.

Letting $\mu \to 0$

$$\lim_{\mu \to 0} x_2(\mu) = -1$$ APPROACHES OPTIMAL OBJECTIVE VALUE

ALGORITHM (Penalty Function)

Given $x_0^s \in \mathbb{R}^n$ and $\mu_0 > 0$. Let $k = 0$.

While $\mu_k > \varepsilon$

(1) Use Newton's method to find an approximate local minimizer x_k of $P(x, \mu_k)$ w.r.t. x such that

$$\| \nabla_x P(x_k, \mu_k) \| < \tau_k .$$

(When choosing the initial guess x_k^s for Newton's method use the previous solution x_{k-1}.)

(2) Choose a $\mu_{k+1} \in (0, \mu_k)$

(3) $k = k+1$

end

THM (Convergence of Penalty Function Method)

Suppose

$$\lim_{k \to \infty} \mu_k = 0 \quad \text{and} \quad \lim_{k \to \infty} \tau_k = 0.$$

Then

(i) any limit point x_* of the sequence $\{x_k\}$ where LICQ holds satisfies the KKT conditions for (NEP) for some Lagrange multiplier λ.

(ii) Furthermore

$$\lim_{k\to\infty} -\frac{\tilde{c}_j(x_k)}{\mu_k} = (\lambda_*)_j .$$

PROOF

Consider

$$\nabla_x P(x_k, \mu_k) = \nabla f(x_k) + \frac{1}{\cancel{2}\mu_k} \sum_{j=1}^{\ell} \tilde{c}_j(x_k) \nabla \tilde{c}_j(x_k)$$

Since $\|\nabla_x P(x_k, \mu_k)\| < \tau_k$ and $\tau_k \to 0$ as $k \to \infty$, we have

$$(*) \quad \lim_{k\to\infty} \left\| \nabla f(x_k) + \sum_{j=1}^{\ell} \frac{\tilde{c}_j(x_k)}{\mu_k} \nabla \tilde{c}_j(x_k) \right\| = 0$$

$$\Longrightarrow$$

$$\lim_{k\to\infty} \left\| \sum_{j=1}^{\ell} \frac{\tilde{c}_j(x_k)}{\mu_k} \nabla \tilde{c}_j(x_k) \right\| = \lim_{k\to\infty} \|\nabla f(x_k)\|$$

$$\Longrightarrow$$

$$\left\| \sum_{j=1}^{\ell} \tilde{c}_j(x_*) \nabla \tilde{c}_j(x_*) \right\| = \lim_{k\to\infty} \mu_k \|\nabla f(x_k)\|$$

$$= 0$$

By assumption $\{\nabla \tilde{c}_j(x_*) : j = 1, \dots, \ell\}$ is linearly independent meaning

$$(1) \quad \tilde{c}_j(x_*) = 0 \qquad j = 1, \dots, \ell .$$

Define

$$\lambda_k := -\frac{\tilde{c}(x_k)}{\mu_k},$$

and note that (*) implies

$$\lim_{k\to\infty} \nabla f(x_k) - J(x_k)^T \lambda_k = 0$$

$$\Longrightarrow$$

$$(2) \quad \nabla f(x_*) = J(x_*)\, \lambda_*$$

where $\lambda_* = \lim_{k\to\infty} \lambda_k$. Equations (1) and (2) show that KKT conditions hold at x_* for (NEP).

$\square$