constrained nonlinear programming

goal: sufficient conditions for local optimality in NLP subject to constraints

refs:

OPTIMAL CONTROL AND ESTIMATION

Robert F. Stengel
Department of Mechanical and Aerospace Engineering
Princeton University, Princeton, New Jersey

Section 2.1

Nonlinear Programming

SECOND EDITION

Dimitri P. Bertsekas
Massachusetts Institute of Technology

Chapter 3

- consider NLP with constraints:

$$\min_{w \in \mathbb{R}^\ell} c(w) \qquad c : \mathbb{R}^\ell \to \mathbb{R}$$
$$\text{s.t. } f(w) = 0 \qquad f : \mathbb{R}^\ell \to \mathbb{R}^n$$

→ assuming you can explicitly solve the constraint equation show how you can reduce to an equivalent unconstrained NLP

- suppose $w \in \mathbb{R}^\ell$ can be "split" into two pieces, $w = (x, u) \in \mathbb{R}^n \times \mathbb{R}^m$ and explicitly solve $f(x,u) = 0$ for x in terms of u,

i.e. $x : \mathbb{R}^m \to \mathbb{R}^n$ $\rightsquigarrow$ then $\min_{u \in \mathbb{R}^m} c(x(u), u)$ is equivalent to original problem

$: u \mapsto x(u)$

i.e. $X : \mathbb{R}^m \to \mathbb{R}^n$ $\rightsquigarrow$ then $\min_{u \in \mathbb{R}^m} c(X(u), u)$ is equivalent
$: u \mapsto X(u)$ to original problem

- can't solve $f(w) = 0$ in general, so we'll consider another approach

<u>idea</u>: consider augmented cost function

$$\tilde{c}(x, u, \lambda) = c(x, u) + \lambda \cdot f(x, u)$$

$\lambda \in \mathbb{R}^{1 \times n}$ termed "<u>Lagrange multipliers</u>" or "<u>dual variables</u>"

and find stationary points of $\tilde{c}$

i.e. want $\underbrace{(x_0, u_0, \lambda_0)}_{\in \mathbb{R}^{(n+m+n)}}$ s.t. $\underbrace{D\tilde{c}(x_0, u_0, \lambda_0)}_{\in \mathbb{R}^{1 \times (n+m+n)}} = 0$

→ determine necessary condition for λ_0 to be stationary (assume $D_x f(x_0, u_0) \in \mathbb{R}^{n \times n}$ invertible / nonsingular)

- $D_x \tilde{c} = D_x c + \lambda \cdot D_x f$ so if $D_x f(x_0, u_0)$ invertible then it's necessary that $\lambda_0 = -D_x c(x_0, u_0) \cdot [D_x f(x_0, u_0)]^{-1}$

<u>def</u>: $(x_0, u_0) \in \mathbb{R}^{n+m}$ is a <u>stationary point</u> for $\min_{(x,u) \in \mathbb{R}^{n+m}} c(x, u)$
s.t. $f(x, u) = 0 \in \mathbb{R}^n$

if $D_u c(x_0, u_0) + \lambda_0 \cdot D_u f(x_0, u_0) = 0$ and $f(x_0, u_0) = 0$
where $\lambda_0 = -D_x c(x_0, u_0) \cdot [D_x f(x_0, u_0)]^{-1}$

→ why is it reasonable to assume $D_x f(x_0, u_0)$ invertible?

- otherwise constraints are redundant!

- otherwise constraints are redundant!

  consider the linear case: $f(x,u) = L \cdot \begin{bmatrix} x \\ u \end{bmatrix}$

$\rightarrow$ why are stationary points of constrained NLP the same as those of $\tilde{c}$?

$*$ full answer requires some work — cf wikipedia article "Lagrange multipliers"

$\rightsquigarrow$ when $n=1$ the geometry is easier to understand — cf pg 103 of Folland's Advanced Calculus

ex: (Lewis Syrmos 1.2-3) quadratic cost, linear constraint:

$$\min\ c(x,u) = \frac{1}{2} x^T Q x + \frac{1}{2} u^T R u \qquad \text{— assume } Q^T = Q > 0,\ R^T = R > 0$$

$$\text{s.t. } f(x,u) = x + Bu + b = 0$$

$\rightarrow$ write augmented cost & stationarity conditions

- $\tilde{c}(x,u,\lambda) = c(x,u) + \lambda \cdot f(x,u)$

$$= \frac{1}{2} x^T Q x + \frac{1}{2} u^T R u + \lambda \cdot (x + Bu + b)$$

- $D_x \tilde{c} = D_x c + \lambda \cdot D_x f = x^T Q + \lambda = 0$
- $D_u \tilde{c} = D_u c + \lambda \cdot D_u f = u^T R + \lambda \cdot B = 0$
- $D_x \tilde{c} = f(x,u) = x + Bu + b = 0$

$\rightarrow$ solve for stationary u (& λ)

- from $D_x \tilde{c} = 0$ we find $\lambda = -x^T Q \iff \lambda^T = -Qx$

- from $D_x C = 0$ we find λ [illegible]
- from $D_u \tilde{C} = 0$ we find $u = -R^{-1} B^T \lambda^T$
- combining these we find $u^* = -R^{-1} B^T Q x$