constrained nlp

Friday, December 23, 2016 12:59 PM

Stengel pg29–41

Lewis et al Ch 1

Bertsekas Ch 3

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

o now consider constrained NLP:

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

$c : \mathbb{R}^\ell \to \mathbb{R}$, $f : \mathbb{R}^\ell \to \mathbb{R}^n$

→ show how (in principle) to reduce this to unconstrained NLP

(1–2 min)

– if we split $w \in \mathbb{R}^\ell$ into $w = (x, u) \in \mathbb{R}^n \times \mathbb{R}^m$,

solve equation $f(x,u) = 0$ for x in terms of u, $x : \mathbb{R}^m \to \mathbb{R}^n$, then could equivalently solve

$$\min_{u \in \mathbb{R}^m} c(x(u), u)$$

- in general, can't solve $f(x,u) = 0$ for x
  - instead, augment objective function:

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

    so now there are $n + m + n$ unknowns:

    $$x \in \mathbb{R}^n, \; u \in \mathbb{R}^m, \; \lambda \in \mathbb{R}^{1 \times n}$$

  - for (x_0, u_0, λ_0) to be local min for $\tilde{c}$, it's necessary that $D\tilde{c}(x_0, u_0, \lambda_0) = 0$

    note: $D\tilde{c} = (D_x \tilde{c}, D_u \tilde{c}, D_\lambda \tilde{c}) \in \mathbb{R}^{n+m+n}$

→ determine nec. cond. on λ_0, assuming $D_x f(x_0, u_0)$ invertible (2-3 min)

- $D_x \tilde{c} = D_x c + \lambda D_x f$, so if $D_x f(x_0, u_0)$ invertible then it's nec. that $\lambda_0 = -D_x c(x_0, u_0) [D_x f(x_0, u_0)]^{-1}$
- $D_u \tilde{c} = D_u c + \lambda D_u f$, so nec. that $D_u c(x_0, u_0) + \lambda_0 D_u f(x_0, u_0) = 0$

- $D_\lambda \tilde{c} = f$, so also nec. that $f(x_0, u_0) = 0$

- we now have enough equations to specify stationary points for constrained NLP:

def: $(x_0, u_0) \in \mathbb{R}^{n \times m}$ <u>stationary point</u> for

$$\min_{(x,u) \in \mathbb{R}^{n \times m}} c(x,u) \quad \text{s.t.} \quad f(x,u) = 0$$

if $D_u c(x_0, u_0) + \lambda_0 D_u f(x_0, u_0) = 0$

$f(x_0, u_0) = 0$, $D_x f(x_0, u_0)$ invertible,

$$\lambda_0 = -D_x c(x_0, u_0) \left[D_x f(x_0, u_0) \right]^{-1}$$

→ why is it reasonable to assume $D_x f$ invertible?

- otherwise constraints are redundant (consider linear case $f(x,u) = L \cdot \begin{bmatrix} x \\ u \end{bmatrix}$)

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

- (this actually requires some work...

cf wikipedia article on Lagrange multipliers)

- the geometry is easy to understand when $n=1$: (taken from pg 103 in Folland's Advanced Calc.)

  - suppose c has a local min at (x_0, u_0)
  - if $\gamma : (-1,1) \to \{(x,u) : f(x,u) = 0\}$ is C^1 and $\gamma(0) = (x_0, u_0)$, then $c \circ \gamma : (-1,1) \to \mathbb{R}$ has local min at 0, so $Dc(x_0, u_0) \cdot D\gamma(0) = 0$
  - this implies $Dc(x_0, u_0)$ is orthogonal to $\{(x,u) : f(x,u) = 0\}$ at (x_0, u_0)
  - since $Df(x_0, u_0)$ is also orthogonal to this submanifold, $Dc(x_0, u_0) = \lambda \cdot Df(x_0, u_0)$, i.e. local mins are stationary points

→ what does it mean if $D_u f$ has full row rank?

- it means we could solve for u in terms of x... see this example:

<u>ex</u> (Lewis Syrmos 1.2-3):

- quadratic cost, linear constraint:

$$\min c(x,u) = \frac{1}{2} x^T Q x + \frac{1}{2} u^T R u$$

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

- assume $Q^T = Q > 0$, $R^T = R > 0$

→ write augmented cost, first-order nec. conds for optimality

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

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

1°. $D_x \tilde{c} = D_x c(x,u) + \lambda D_x f(x,u)$

$$= x^T Q + \lambda = 0$$

2°. $D_u \tilde{c} = u^T R + \lambda B = 0$

3°. $D_\lambda \tilde{c} = f(x,u)$

$$= x + Bu + b = 0$$

- due to linearity, we can solve:

1°. to find $\lambda = -Qx$

2°. to find $u = -R^{-1} B^T \lambda^T$ } why is R invertible?

$\rightarrow R$ invertible?

3°: says $\lambda = QBu + Qb$

- thus $u = -R^{-1}B^T(QBu + Qb)$

characterizes stationary control input

$\rightarrow$ solve for u; why can you invert the matrix?

- equivalent to $(R + B^TQB)u = -B^TQb$
- since $R > 0$, $B^TQB \geq 0$,

  $(R + B^TQB) > 0$, so its invertible
- $u = -(R + B^TQB)^{-1}B^TQb$