OPTIMAL CONTROL

JOSÉ C. GEROMEL

DSCE / School of Electrical and Computer Engineering
UNICAMP, 13083-852, Campinas, SP, Brazil,
geromel@dsce.fee.unicamp.br

Campinas, Brazil, July 2012

Contents

Unconstrained problem

- Consider the optimization problem

$$\min_{x \in \mathbb{R}^n} f(x)$$

where $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is a known function.

Definition (Local and global minimum)

A point $x^* \in \mathbb{R}^n$ is a local minimum of $f(\cdot)$ if

$$f(x^*) \leq f(x), \;\; \forall x \in \Omega$$

where Ω is a neighbourhood of x^*. If this inequality remains true for $\Omega \equiv \mathbb{R}^n$ then x^* is a global minimum of $f(\cdot)$.

- Convexity implies that any local minimum is global.

Unconstrained problem

- For $f(\cdot) \in \mathbb{C}^2$, twice continuously differentiable, we define:
- The gradient vector

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1} \cdots \frac{\partial f(x)}{\partial x_n} \right]' \in \mathbb{R}^n$$

- The hessian (symmetric) matrix

$$\nabla^2 f(x) = \left\{ \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \ \forall i,j = 1, \cdots, n \right\} \in \mathbb{R}^{n \times n}$$

Theorem (1)

Whenever $f(\cdot) \in \mathbb{C}^2$ the following hold:

- *If $x^* \in \mathbb{R}^n$ is a local minimum then $\nabla f(x^*) = 0$.*
- *If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) > 0$ then $x^* \in \mathbb{R}^n$ is a local minimum.*

Unconstrained problem

- Proof : The proof is based on Taylor's series development

$$f(x) = f(x^*) + \epsilon \nabla f(x^*)' y + \frac{\epsilon^2}{2} y' \nabla^2 f(x^*) y + \mathcal{O}(\epsilon^3)$$

where $x = x^* + \epsilon y$.

  - To prove the first part, notice that for $y = -\nabla f(x^*)$ there exists $\epsilon > 0$ small enough such that $f(x) - f(x^*) < 0$ unless $\nabla f(x^*) = 0$.
  - For the second part we notice that

$$f(x) > f(x^*) + \mathcal{O}(\epsilon^3)$$

indicating that for $\epsilon > 0$ small enough the claim follows.

Unconstrained problem

- Some important remarks:
  - In the second part of the theorem, we can not replace $\nabla^2 f(x^*) > 0$ by $\nabla^2 f(x^*) \geq 0$. Why?
  - Convexity imposes that
    $$f(x) \geq f(x^*) + \nabla f(x^*)'(x - x^*)$$
    for all $x \in \mathbb{R}^n$. In this case, $\nabla f(x^*) = 0$ is a necessary and sufficient condition for optimality.
  - Convexity can be characterized alternatively by the condition $\nabla^2 f(x) \geq 0, \ \forall x \in \mathbb{R}^n$. Why?

Constrained problem

- Now, consider the optimization problem

$$\min_{x \in \mathbb{R}^n} \{ f(x) \ : \ g(x) = 0 \}$$

where $g(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$. It is assumed that $m \leq n$ and the feasible set defined by all $x \in \mathbb{R}^n$ such that $g(x) \leq 0$ is not empty.

Definition (Constrained minimum)

A point $x^* \in \mathbb{R}^n$ is a local constrained minimum if $g(x^*) = 0$ and

$$f(x^*) \leq f(x), \ \forall x \in \{g(x) = 0\} \cap \Omega$$

If $\Omega \equiv \mathbb{R}^n$ it is a global constrained minimum.

Constrained problem

- Optimality can be characterized by the so called Lagrangian

$$L(x, p) = f(x) + p'g(x)$$

where $p \in \mathbb{R}^m$ is called Lagrange multiplier. Since $L(x, p) = f(x)$ for all $x \in \mathbb{R}^n$ such that $g(x) = 0$, the next result follows immediately from Theorem (1).

Theorem (2)

Whenever $f(\cdot) \in \mathbb{C}^1$ and $g(\cdot) \in \mathbb{C}^1$, if $x^* \in \mathbb{R}^n$ is a local constrained minimum then

$$\frac{\partial L(x^*, p^*)}{\partial x} = 0 \ , \ \frac{\partial L(x^*, p^*)}{\partial p} = 0$$

Constrained problem

- It is important to observe that :
  - Theorem (2) is only a necessary condition for optimality. In general, only for convex problems sufficiency also holds.
  - A convex problem of this type is characterized by having $f(x)$ convex and $g(x)$ linear. Why?
  - Notice that

$$\frac{\partial L(x^*, p^*)}{\partial p} = g(x^*) = 0$$

    hence, x^* is always feasible. This is the key property of the Lagrangian.

Constrained problem

- The next theorem provides a sufficient condition for optimality.

Theorem (3)

Let $p^* \in \mathbb{R}^m$. If $x^* \in \mathbb{R}^n$ is such that $g(x^*) = 0$ and solve

$$\min_{x \in \mathbb{R}^n} L(x, p^*)$$

then $x^* \in \mathbb{R}^n$ is a global constrained minimum.

Proof : It is remarkably simple. It follows from the assumption $g(x^*) = 0$ and $L(x^*, p^*) \leq L(x, p^*), \ \forall x \in \mathbb{R}^n$

$$\underbrace{f(x^*) + (p^*)' g(x^*)}_{=f(x^*)} \leq \underbrace{f(x) + (p^*)' g(x), \ \ \forall x \in \mathbb{R}^n}_{=f(x), \ \forall g(x)=0}$$

Constrained problem

- Some important remarks:
  - If the minimization in Theorem (3) is replaced by

    $$\frac{\partial L(x^*, p^*)}{\partial x} = 0$$

    then the necessary condition in Theorem (2) is readily obtained.
  - For convex problems with $f(x)$ convex and $g(x)$ linear both theorems provide necessary and sufficient conditions for optimality.

Special constrained problem

- We want to put in evidence an optimization problem that presents the same structure of the optimal control problems

$$\min_{u \in U} \{ f(x,u) \ : \ x = G(u) \}$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. The set $U \subset R^m$ imposes constraints on the variable $u \in \mathbb{R}^m$ exclusively.

- The Lagrangian associated to this problem is

$$L(x,u,p) = f(x,u) + p'(G(u) - x)$$

which from Theorem (3) yields the next optimality characterization.

Special constrained problem

- The next theorem provides sufficient conditions for optimality.

Theorem (4)

Let $p^* \in \mathbb{R}^n$. If the pair (x^*, u^*) is such that $x^* = G(u^*)$ and solve

$$\min_{x,u \in U} L(x, u, p^*)$$

then (x^*, u^*) is a global constrained minimum.

Proof : It is left as an exercise.

- Alternatively, from the definition of the Hamiltonian as

$$H(x, u, p) = f(x, u) + p' G(u)$$

optimality is characterized as well.

Special constrained problem

- The relationship between the Lagrangian and Hamiltonian

$$L(x, u, p) = H(x, u, p) - p'x$$

allows us to rewrite conditions of Theorem (4) as:

1. $G(u^*) = x^*$
2. $\min_x H(x, u^*, p^*) - (p^*)'x$
3. $\min_{u \in U} H(x^*, u, p^*)$

Moreover, since there is no further constraint on $x \in \mathbb{R}^n$, if Theorem (2) is replaced by its necessary condition of optimality we have:

1. $\frac{\partial H}{\partial p}(x^*, u^*, p^*) = x^*$
2. $\frac{\partial H}{\partial x}(x^*, u^*, p^*) = p^*$
3. $\min_{u \in U} H(x^*, u, p^*)$

which leads to the forthcoming important result.

Special constrained problem

- The next theorem provides necessary conditions for optimality.

Theorem (5)

If the pair (x^*, u^*) is a local constrained minimum then:

- u^* *solves* $\min_{u \in U} H(x^*, u, p^*)$
- $\frac{\partial H}{\partial x}(x^*, u^*, p^*) = p^*$
- $\frac{\partial H}{\partial p}(x^*, u^*, p^*) = x^*$

Proof : Immediate from the previous discussion.

- We observe that all optimality conditions are expressed, exclusively, in terms of the Hamiltonian associated to the problem under consideration.

Special constrained problem

- We have to add the following remarks:
  - Function $H(x, u, p)$ does not need to be differentiable with respect to $u \in \mathbb{R}^m$. However, if it is differentiable and $U \equiv R^m$ the first condition can be replaced by

    $$\frac{\partial H}{\partial u}(x^*, u^*, p^*) = 0$$

    with no loss of generality.
  - The result becomes sufficient if $x^* \in \mathbb{R}^n$ solves

    $$\min_x H(x, u^*, p^*) - (p^*)'x$$

    This occurs whenever $f(x, u)$ is convex with respect to $x \in \mathbb{R}^n$ for each $u \in U$.

Special constrained problem

- The problem under consideration can be rewritten as

$$\min_{u \in U} J(u)$$

where $J(u) = f(\underbrace{x}_{=G(u)}, u)$. Using the chain rule we have

$$\begin{aligned} \frac{dJ}{du}(u) &= \frac{\partial f}{\partial u}(x,u) + \underbrace{\frac{\partial f}{\partial x}(x,u)'}_{p'} \frac{\partial G}{\partial u}(u) \\ &= \frac{\partial H}{\partial u}(x,u,p) \end{aligned}$$

The gradient of $J(u)$ is equal to the partial derivative of the Hamiltonian function! This explains the previous results.

Suggested problems

- **Problem** 1: Let $A(t) : \mathbb{R} \to \mathbb{R}^{n \times n}$. Determine

$$\frac{d}{dt}A(t)^{-1}, \ \frac{d}{dt}\mathrm{tr}\{A(t)^{-1}\} \ , \ \frac{d}{dt}\det\{A(t)^{-1}\}$$

- **Problem** 2: Let $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$ and

$$H = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}$$

Show that for all $t \geq 0$ and the same partition

$$e^{Ht} = \begin{bmatrix} e^{At} & \int_0^t e^{A\tau} B d\tau \\ 0 & I \end{bmatrix}$$

Suggested problems

- **Problem** 3: Consider $f(x)$ a convex function and $S \subset \mathbb{R}^n$ a polyhedral convex set defined by the extreme points $x_i \in \mathbb{R}^n$ for $i = 1, \cdots, N$. Show that the optimal solution of $\max_{x \in S} f(x)$ is an extreme point of S.
- **Problem** 4: Let $g \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$ of full row rank be given. Solve the problem

$$\min_x \{\|g - x\|_2^2 \;:\; Ax = 0\}$$

where $\|x\|_2^2 = x'x$. What happens if a) $Ag = 0$ and b) there exits $z \in \mathbb{R}^m$ such that $g = A'z$.

Suggested problems

- **Problem** 5: Let $Q \in \mathbb{R}^{n\times n}$ be a symmetric definite positive matrix and $A \in \mathbb{R}^{m\times n}$ of full row rank. Solve the quadratic programming problem

$$\min_x \{x'Qx \;:\; Ax = b\}$$

- **Problem** 6: Let $A \in \mathbb{R}^{m\times n}$ be given. Show that

$$\lambda_m \leq \frac{x'A'Ax}{x'x} \leq \lambda_M, \forall x \neq 0$$

  where λ_m and λ_M are the minimum and the maximum eigenvalue of $A'A$. Based on this result show that $\|A\| = \sqrt{\lambda_M}$ is a valid norm for matrix A.

Suggested problems

- **Problem 7:** Convert the program

$$\min_{x}\{\max_{i=1,\cdots,n} c_i' x \ : \ Ax = b\}$$

to a linear program and calculate its dual.

- **Problem 8:** Consider the scalar function

$$f(\alpha) = \det(A + \alpha B)$$

where A and B are square matrices with A nonsingular. Determine de derivative of $f(\cdot)$ at $\alpha = 0$.

Suggested problems

- **Problem** 9: Consider the function $v(y) : \mathbb{R}^n \to \mathbb{R}$ defined by

$$v(y) = \min_{x \in S} f(x, y)$$

where $f(x, y)$ is linear with respect to $y \in \mathbb{R}^n$. Show that $v(y)$ is concave.

- **Problem** 10: Consider the function $v(y) : \mathbb{R}^n \to \mathbb{R}$

$$v(y) = \min_{x \in S} \{ f(x) \ : \ Ax \leq y \}$$

where $f(x) : \mathbb{R}^m \to \mathbb{R}$. Determine a lower bound to the right directional derivative at $y_0 \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^m$, defined by

$$D_+ v(y_0, d) = \lim_{\varepsilon \to 0^+} \frac{v(y_0 + \varepsilon d) - v(y_0)}{\varepsilon}$$