unconstrained nlp

Stengel pg 29–41

Lewis et al Ch 1

Bertsekas 99 Ch 1

goal: necessary & sufficient conditions
for local optimality in
NonLinear Programs (NLP)

- suppose I ask you to minimize a function
  $c : \mathbb{R}^m \to \mathbb{R}$, i.e. to "solve" the
  $: u \mapsto c(u)$

  "nonlinear program" (NLP) $\min_{u \in \mathbb{R}^m} c(u)$

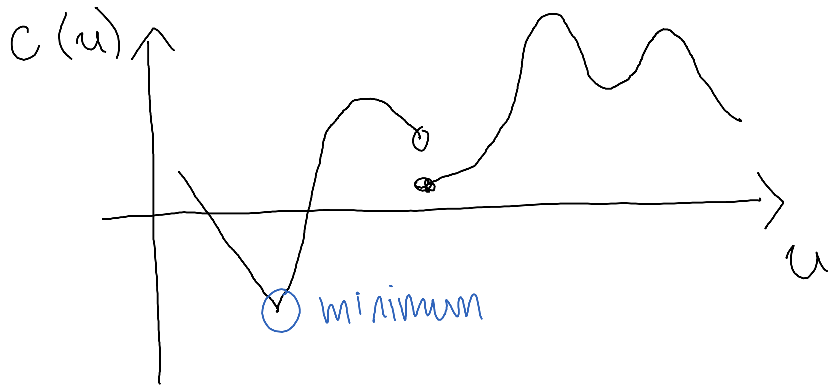
  – c called the <u>objective function</u>

  – note that " $\min_{u \in \mathbb{R}^m} c(u)$ " is just notational
  shorthand for " $\min\{c(u) \mid u \in \mathbb{R}^n\}$ ",
  which is a function (min) applied to a set

→ what would you do when $m=1$?

→ does your strategy generalize to $m>1$?

- when $m=1$, can just look at the graph:



* this "gestalt" solution method might work for $m=2$, but fails for $m>2$ because $\text{graph}(c) \subset \mathbb{R}^{m+1}$, and we can only "look" in 3 dimensions

* note that we can't "look" at all possible u's since there's an infinite # of u's
  - motivates looking "locally":

def: $u \in \mathbb{R}^m$ is a <u>local minimizer</u> for (NLP) if there exists an open set $V \subset \mathbb{R}^m$ containing u such that

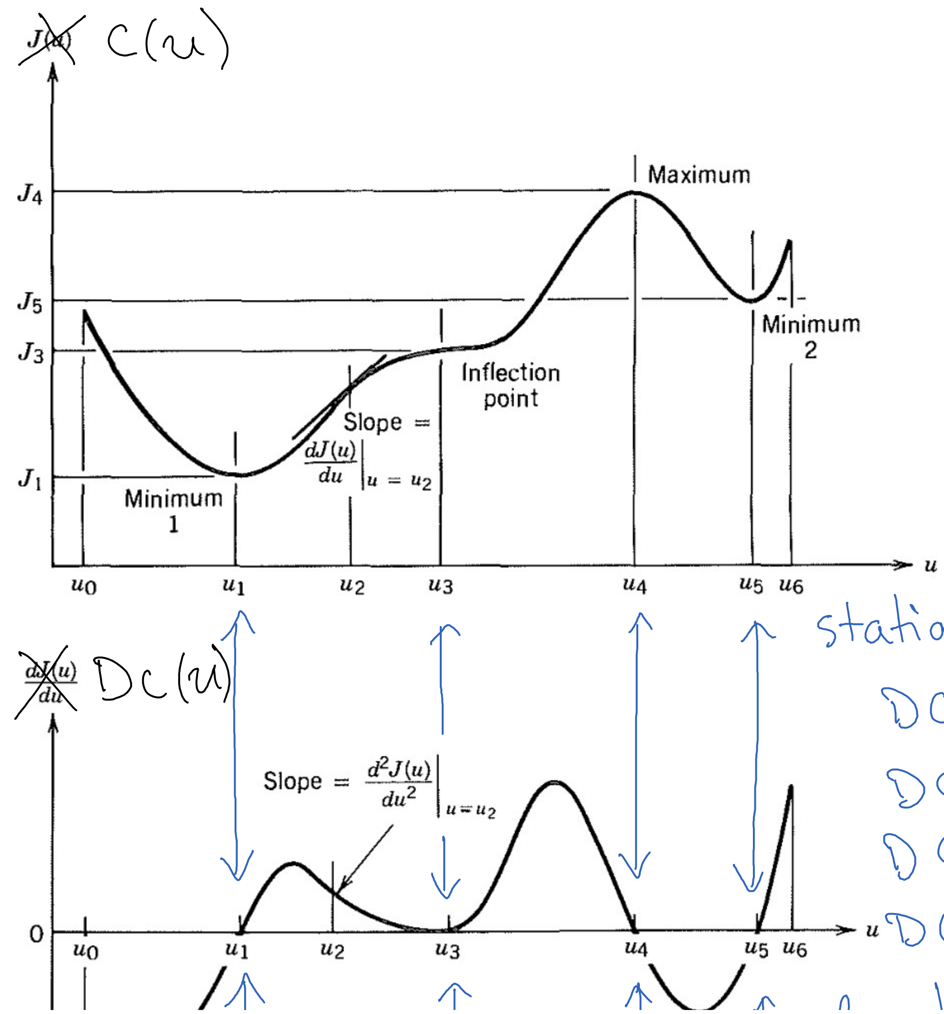
$\forall v \in V: c(u) \leq c(v)$

- strict local min if inequality strict ($<$)
- u is a local maximizer if it's a local min for $\min_{u \in \mathbb{R}^m} -c(u)$

* difficult to directly operationalize, since it requires checking inequality holds at $|V| = \infty$ points

∘ a more systematic approach uses derivatives:

FIGURE 2.1-1 Scalar function of a scalar variable, showing locally minimum, stationary, and maximum values.



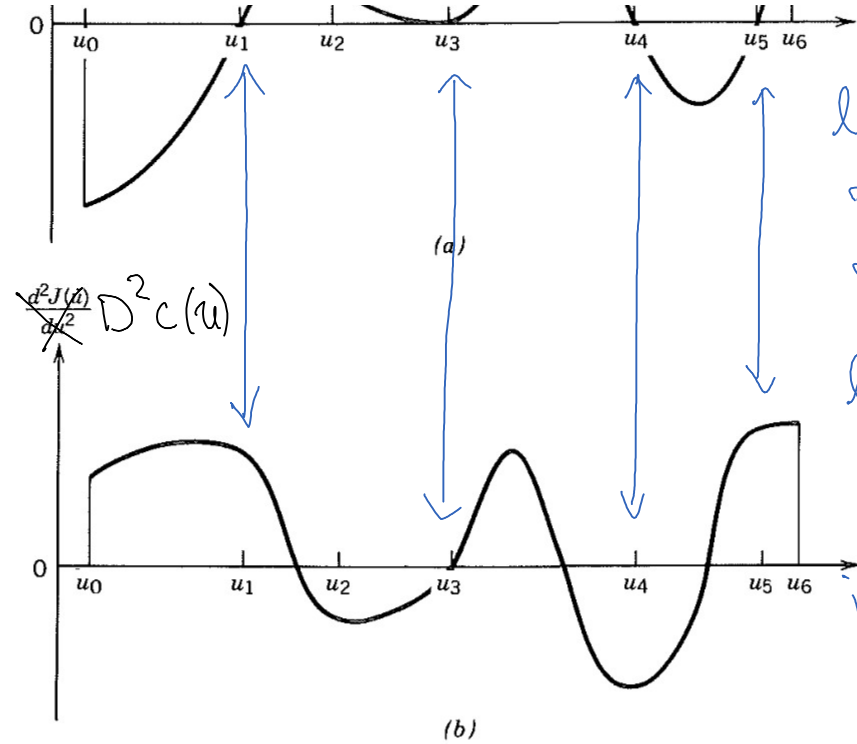
stationary points:

$Dc(u_1) = 0$

$Dc(u_4) = 0$

$Dc(u_3) = 0$

$Dc(u_5) = 0$



$Dc(u_5) = 0$

local min:

$D^2c(u_1) > 0$

$D^2c(u_5) > 0$

local max:

$D^2c(u_4) < 0$

indeterminant:

$D^2c(u_3) = 0$

FIGURE 2.1-2 Slopes and curvatures of $J(u)$, $dJ(u)/du$, and $d^2J(u)/du^2$. (*a*) First derivative of $J(u)$, $dJ(u)/du$; (*b*) second derivative of $J(u)$, $d^2J(u)/du^2$.

– this approach generalizes to $m > 1$:

<u>def</u>: $u_0 \in \mathbb{R}^m$ is a <u>stationary point</u> for (NLP) if $Dc(u_0) = 0$

*assumes c is C^1 at u_0

<u>thm</u>: (sufficient conditions for optimality)

a stationary point $u_0 \in \mathbb{R}^m$ is:

– a strict local min if $D^2c(u_0) > 0$

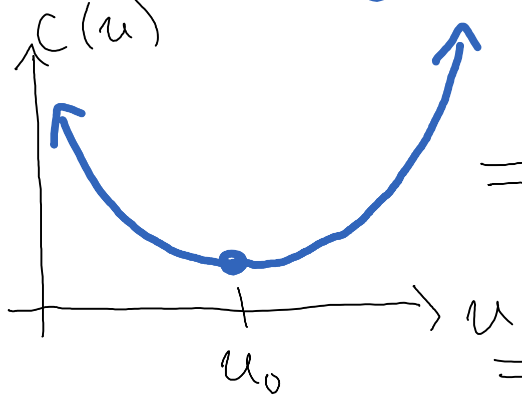
– a strict local max if $D^2c(u_0) < 0$

*assumes c is C^2 at u_0

* assumes c is C^2 at u_0

$\Rightarrow D^2c(u_0)$ is symmetric, so makes sense

ex:



$c(u) = u^2$

$\Rightarrow Dc(0) = 0$

$D^2c(0) = 2 > 0$

$\Rightarrow u_0 = 0$ is local min

$\rightarrow$ determine c s.t. $u_0 = 0$ is local min but $\cancel{D^2c(u_0) > 0}$

note: not C^1

$- c(u) = 0 \quad - c(u) = u^4 \quad - c(u) = |u|$

thm (necessary conditions for optimality):

if $u_0 \in \mathbb{R}^m$ is a local min for (NLP):

- if c is C^1 at u_0 then $Dc(u_0) = 0$
- if c is C^2 at u_0 then $D^2c(u_0) \geq 0$

ex: suppose 2nd-order approx to c at $u_0 \in \mathbb{R}^m$ is exact, i.e.:

$C \in \mathbb{R}^{m \times m}$

$$(c(u) - c(u_0)) = b^T(u - u_0) + \frac{1}{2}(u - u_0)^T C (u - u_0):$$

$\neq c : \mathbb{R}^m \to \mathbb{R}$

→ determine necessary conditions on $b^T \in \mathbb{R}^{1\times m}$, $C^T = C$ for $u_0 \in \mathbb{R}^m$ to be local min

→ determine sufficient cond's on spec $D^2c(v)$ for $u_0 \in \mathbb{R}^m$ to be strict local min

→ if strict local min, solve for u_0

- necessary that $Dc(u_0) = b^T + (u_0 - u_0)^T C$

$$= b^T = 0$$

and $D^2 c(u_0) = C \geq 0$