Math 3 April 23rd 2019 - last lecture

... unless... an "exam review" session?

YES. Trying to schedule one May 16th

Topics:

* How to easily destroy Lagrange (/K-T)

* How to deal properly with that issue?

And to that end: why does Lagrange work?

* Lagrange as an application of [failure of] the implicit function theorem

  ↳ $g \in C^1$ not sufficient for smooth level curves!

  (But if $\nabla g \neq \vec{0}^T$ .... it is.)

An application:

* Risk / return trade-off: why textbooks rather minimize variance for given expected return, than max E[return] given variance.

And a leftover differential equation of first order:

* Bernoulli's differential equation

Destroying Lagrange

The set S given by $g(\vec{x}) = b$ is also given by $(g(\vec{x}) - b)^3 = 0$.

Lagrangians for $\max/\min_{\vec{x} \in S} f(\vec{x})$:

$$f(\vec{x}) - \lambda(g(\vec{x}) - b) \quad \text{or alternatively}$$

$$f(\vec{x}) - \mu(g(\vec{x}) - b)^3$$

FOC: $$\nabla f(\vec{x}) = \begin{cases} \lambda \, \nabla g(\vec{x}) \\ \underbrace{3(g(\vec{x}) - b)^2}_{=0} \mu \, \nabla g(\vec{x}) \end{cases}$$

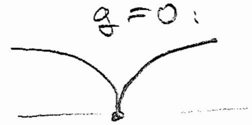
If $\vec{x}^*$ optimal and $\nabla f(\vec{x}^*) \neq \vec{0}^T$, then the "$\lambda$" version is OK, but there is no μ such that $3(\underbrace{g(\vec{x}^*) - b}_{=0})^2 \mu = \lambda \neq 0$.

The problem is: points where $\nabla[(g-b)^3] = \vec{0}^T$

- In this case: could have been avoided.
- Sometimes it cannot, if you want ∇g to exist at all.

Example: $g = y^3 - x^2$

$g=0$:



Necessary Kuhn-Tucker cond's, C^1 case.

Problem: $\max f(\vec{x})$ s.t. $g_j(\vec{x}) \le b_j$ some j's
$g_j(\vec{x}) = b_j$ some j's

Lagrangian: $f(\vec{x}) - \sum \lambda_j (g_j(\vec{x}) - b_j)$

- Consider $\sum \lambda_j \nabla g_j(\vec{x})$
- Delete the terms for which $g_j(\vec{x}) < b_j$.
- Remains: a linear combination of $\nabla g_j(\vec{x})$ (row vectors).

  If those are linearly dependent, there could be a max there even if the "Math 2 K-T cond's" fail.

So, conditions: Let $\vec{x}^*$ solve the problem.

Then there exist $\lambda_1, \dots, \lambda_m$, $\lambda_j = 0$ if $g_j(\vec{x}^*) < b_j$,
and $\lambda_j \ge 0$ for the ineq. constraints,
such that

$\sum \lambda_j \nabla g(x^*)$ is $\begin{cases} \text{either } = \nabla f(\vec{x}^*) \\ \text{or } \left| \begin{array}{l} = \vec{0}^T, \text{ and not} \\ \text{all the } \lambda_j \text{ are zero} \end{array} \right. \end{cases}$

The conditions that the gradients of the g_i are linearly independent is called a constraint qualification ("CQ").

- If CQ holds at $\vec{x}^*$ then the K-T cond's are truly necessary.
- In other words: At optimum, either K-T hold or CQ fails.

- There is a "unified version": ("Fritz John")

  Let $L(\vec{x}) = \lambda_0 f(\vec{x}) - \sum \lambda_j (g_j(\vec{x}) - b)$

  If $\vec{x}^*$ optimal, there exist $\lambda_0, \ldots, \lambda_m$, not all $= 0$, and with $\lambda_0 = 0$ or 1

  such that

  $\nabla L(\vec{x}^*) = \vec{0}^T$

  $\lambda_j \geq 0$ ($= 0$ if $g_j(\vec{x}^*) < b_j$) for $\leq$ constraints

  $g_j(\vec{x}^*)$ for $=$ constraints.

Note: If the admissible set ends
in a "tip", e.g. 
you ___cannot___ expect the Math 2
cond's to produce the tip point
even if it is optimal!

What do you need for the (2019) exam?

* To know that it "isn't as simple as Math2"
* "Worst" that has ever been asked in a 4140 exam, asked how could there be no point satisfying K-T even if the problem actually had a solution.

  Sufficient answer:
  "Constraint qualification fails at the optimal point."

* You can totally disregard the "concave programs CQ" (the "Slater" condition)
* ... and ... linear constraints ... disregard.

Implicit functions revisited.

Two variables: $g(x, y) = b$, $g \in C'$

* $\frac{dy}{dx} = -\frac{g'_x}{g'_y}$ is continuous as long as $g'_y \neq 0$.

* If $g'_y = 0$, then $\frac{dx}{dy} = -\frac{g'_y}{g'_x}$ .....

$\Rightarrow$ Smooth curve unless $g'_x = g'_y = 0$.

* At stationary points: not necessarily smooth. $y^3 - x^2 = 0$.

* But if rank$(\nabla g) = 1$, then one variable is a C' function of the other – "others" if $n > 2$.

Consider now n variables, $n \geq m+1$.

$g_0(\vec{x}) = b_0$ this g_0 will become $-f$.

$g_1(\vec{x}) = b_1$ .... $g_m(\vec{x}) = b_m$

As long as the Jacobian has <u>full rank</u>

(i.e "full row rank") then $n-(m+1)$ of the x_i are C' functions of the others and the b_j.

- Again: full rank Jacobian $\Leftrightarrow$ cont. function ...
  ... $\Leftrightarrow$ we can change $b_0, \dots, b_m$ a little bit and still find $\vec{x}$ ....

  Let now $g_0 = -f$.
- At max for f, we <u>can no longer</u> reduce b_0; it is the <u>smallest possible</u> given the other b_j.
- So at optimum, the Jacobian of $\begin{pmatrix} -f \\ g_1 \\ \vdots \\ g_m \end{pmatrix}$ cannot have full (row) rank.
- There must be linear dependence between the rows:

$$\lambda_0 \nabla f - \lambda_1 \nabla g_1 - \dots - \lambda_m \nabla g_m = 0$$

for some numbers not all $= 0$.

↳ either λ_0 must be chosen $= 0$ $\Rightarrow$ CQ fails

↳ or λ_0 can be $\neq 0$: then λ_0 can be $= 1$
$\Rightarrow$ K-T FOC.

$|K| \neq 0$
(why assume that?)

* Risk/return trade-off example

- Agent chooses to invest $\vec{v}$ in n risky opportunities, covariance matrix K and mean vector $\vec{r}$; the rest of endowment, $1 - \vec{v}^T \vec{1}$ is invested safe at return r_0 ($r_0 = 0$ for simplicity).

- Provided the returns vector is elliptically distributed – a class including multinormals – everyone who maximizes expected increasing utility will do the following:

$$\max_{\vec{v}} \vec{v}^T \vec{r} \quad \text{s.t} \quad \vec{v}^T K \vec{v} = Q^2 \quad (= \text{variance})$$

K-T: $\vec{r} = 2\lambda K \vec{v}$ and $\Rightarrow \ldots$

- Often you rather see the problem of minimizing variance given mean:

$$\min_{\vec{v}} \vec{v}^T K \vec{v} \quad \text{s.t} \quad \vec{v}^T \vec{r} = R$$

$$2K\vec{v} = \xi \vec{r} \quad \text{and} \Rightarrow \ldots$$

Why do this? It assumes risk aversion, so there must be something appealing .... ?!

Now assume there is <u>not</u> any safe opportunity: that is, impose also $\vec{v}^T\vec{1} = \eta$

$$\vec{r} = 2\lambda \bar{K}\vec{v} + \phi\vec{1}$$

vs for risk-averse:

$$2\bar{K}\vec{v} = \zeta\vec{r} + \psi\vec{1}$$

Any more reason why the latter is "more convenient"?

(Answers: when the plane $\vec{v}^T\vec{1} = \eta$ is <u>tangent</u> to the ellipsoid $\vec{v}^T\bar{K}\vec{v} = Q^2$, [this is the so-called minimum variance portfolio], the CQ fails.

Whereas, the "risk-averse problem" is a convex minimization program

$\leftrightarrow$ "concave program" in Math 3 lingo)

Bernoulli's differential equation

$$\dot{x} + a(t)\,x = b(t)\,x^r$$

r a constant.

We assume $r \neq 0$ (linear, → Math 2)

and $r \neq 1$ (homogeneous linear)

Look for solutions $x > 0$, so x^r defined.

Larger picture: transforming variables can lead to "easier" diff. eqs.

Method: Let $z = x^{1-r}$

Then $\dot{z} = (1-r)\,x^{-r}\,\dot{x}$

which $= (1-r)\,x^{-r}\cdot\left[b(t)\,x^r - a(t)\,x\right]$

$= (1-r)\left[b(t) - a(t)\,\underbrace{x^{1-r}}_{=z}\right]$

So

$$\dot{z} + (1-r)\,a(t)\,z = (1-r)\,b(t).$$

Linear! Solve this. Then $x = z^{\frac{1}{1-r}}$

(Possible to remember: Replace

a by $(1-r)\,a$

b by $(1-r)\,b$

and then replace r by 0.)

Example: Logistic growth, ρ constant:

$$\dot{x} = \rho x \cdot (1 - x/k).$$

(... is separable if k constant, can solve by

$$\frac{k\,dx}{x(k-x)} = \rho\,dt, \quad \frac{k}{x(k-x)} = \frac{1}{x} + \frac{1}{k-x}$$)

But can also be solved as Bernoulli with $r=2$;

$1-r=-1$, leads to

$$\dot{z} = -\rho z + \rho/k$$ (even when k is not constant!)

$$(e^{\rho t} z)^{\cdot} = \rho\, e^{\rho t}/k$$

$$e^{\rho t} \underbrace{z}_{=1/x} = \underbrace{z_0}_{=1/x_0} + B(t), \text{ where } B = \rho \int_0^t \frac{e^{\rho s}}{k(s)}\,ds$$

$$x(t) = \frac{e^{\rho t}}{\frac{1}{x_0} + B(t)} = \frac{x_0 e^{\rho t}}{1 + \rho x_0 \int_0^t \frac{e^{\rho s}}{k(s)}\,ds}.$$

Note: • if $\rho > 0$ and $k(t) \to \bar{k} \in (0,\infty)$,

$$\lim_{t\to+\infty} x(t) = \lim \frac{\rho x_0 e^{\rho t}}{\rho x_0 e^{\rho t}/k(t)} = \lim k(t) = \bar{k}.$$

• Since $r=2$, $z = x^{-1}$ is well-defined even when $x<0$.

But how can $x \to \bar{k}$ then? If x starts negative?

Note, z will then cross zero! Caveat!