Lecture 18: Critical points and optimization

November 16, 2010

Announcements
▶ Final exam dates:
   ▶ Section 7: December 21, 1–4 PM (Note change)
   ▶ Section 8: December 16, 1–4 PM
▶ Pick up midterm, HW 8.
▶ HW 9 due Tuesday, Nov. 23
▶ Regular office hours: Wednesday, Nov. 17

Midterm
You did very well on a challenging and long exam.
▶ Percentiles:
   ▶ High score: 75
   ▶ 75%: 68.5
   ▶ 50%: 64 (median)
   ▶ 25%: 59.5
▶ Approximate grade ranges:
   ▶ 64–75: A range
   ▶ 52–63: B range

One terminology note: A real number is also a complex number.
You did very well on a challenging and long exam.

- Percentiles:
  - High score: 75
  - 75%: 68.5
  - 50%: 64 (median)
  - 25%: 59.5
- Approximate grade ranges:
  - 64–75: A range
  - 52–63: B range

One terminology note: A real number is also a complex number.

Lecture 18: Critical points and optimization

Optimization in general

**Question**
Given a function of two variables, when does it reach its “best” value (minimum or maximum)?
This is extremely important in many contexts.
It is why Economics requires this class!

**Definition**
Point \((a, b)\) is global maximum of \(f(\cdot, \cdot)\) if for any \((x, y)\) \(\in \mathbb{R}^2\), have \(f(a, b) \geq f(x, y)\).
Point \((a, b)\) is local maximum of \(f(\cdot, \cdot)\) if for any \((x, y)\) \(\in \mathbb{R}^2\) close to \((a, b)\), have \(f(a, b) \geq f(x, y)\).
Formally, \(\exists \delta > 0\) so that if \(\| (a - b) - (x, y) \| < \delta\), have \(f(a, b) \geq f(x, y)\).

**Real-world examples**
- What is the biggest box you can make with fixed amount of cardboard?
- What combination of price and quality maximizes revenue?

**Mathematical examples**
What are the maximum/minimum values of
- \(x^2 + y^2\)?
  - Min. at \((0, 0)\). No max.
- \(\cos(\pi x) + \cos(\pi y)\)?
  - Mins and maxes seem to be in grid.
- \(|x| + |y|\)?
  - Min. at \((0, 0)\), but not diff.

**Optimization in general**

**Question**
Given a function of two variables, when does it reach its “best” value (minimum or maximum)?
This is extremely important in many contexts.
It is why Economics requires this class!

**Definition**
Point \((a, b)\) is global maximum of \(f(\cdot, \cdot)\) if for any \((x, y)\) \(\in \mathbb{R}^2\), have \(f(a, b) \geq f(x, y)\).
Point \((a, b)\) is local maximum of \(f(\cdot, \cdot)\) if for any \((x, y)\) \(\in \mathbb{R}^2\) close to \((a, b)\), have \(f(a, b) \geq f(x, y)\).
Formally, \(\exists \delta > 0\) so that if \(\| (a - b) - (x, y) \| < \delta\), have \(f(a, b) \geq f(x, y)\).

**Real-world examples**
- What is the biggest box you can make with fixed amount of cardboard?
- What combination of price and quality maximizes revenue?

**Mathematical examples**
What are the maximum/minimum values of
- \(x^2 + y^2\)?
  - Min. at \((0, 0)\). No max.
- \(\cos(\pi x) + \cos(\pi y)\)?
  - Mins and maxes seem to be in grid.
- \(|x| + |y|\)?
  - Min. at \((0, 0)\), but not diff.
Optimization in general

Question
Given a function of two variables, when does it reach its "best" value (minimum or maximum)?

This is extremely important in many contexts. It is why Economics requires this class!

Definition
Point \((a, b)\) is global maximum of \(f(\cdot, \cdot)\) if for any \((x, y) \in \mathbb{R}^2\), have \(f(a, b) \geq f(x, y)\).

Point \((a, b)\) is local maximum of \(f(\cdot, \cdot)\) if for any \((x, y) \in \mathbb{R}^2\) close to \((a, b)\), have \(f(a, b) \geq f(x, y)\).

Formally, \(\exists \delta > 0\) so that if \(\| (a - b) - (x, y) \| < \delta\), have \(f(a, b) \geq f(x, y)\).

Real-world examples

What is the biggest box you can make with fixed amount of cardboard?

What combination of price and quality maximizes revenue?

Mathematical examples

What are the maximum/minimum values of

\(x^2 + y^2\)

Min. at \((0,0)\). No max.

\(\cos(\pi x) + \cos(\pi y)\)

Mins and maxes seem to be in grid.

\(|x| + |y|\)

Min. at \((0,0)\), but not diff.

1D review

The critical points of \(f(\cdot)\) are the points where \(f'(x) = 0\).

A critical point is

- Local maximum if \(f''(x) < 0\)
- Local minimum if \(f''(x) > 0\)

Maxima and minima occur at critical points, at boundary, or where \(f\) is not differentiable. (Today, we will have no boundary.)

Example

Maximize \(f(x) = x^3 - 3x\) for \(x \in [-3, 3]\).

Solution

- Critical points are where \(f'(x) = 3x^2 - 3 = 0\), or \(x = -1, x = +1\).
- \(f''(1) = 6 > 0\): local min
- \(f''(-1) = -6 < 0\): local max
- Actual maximum: \(f(3) = 18 > 2 = f(-1)\).

1D review

The critical points of \(f(\cdot)\) are the points where \(f'(x) = 0\).

A critical point is

- Local maximum if \(f''(x) < 0\)
- Local minimum if \(f''(x) > 0\)

Maxima and minima occur at critical points, at boundary, or where \(f\) is not differentiable. (Today, we will have no boundary.)

Example

Maximize \(f(x) = x^3 - 3x\) for \(x \in [-3, 3]\).

Solution

- Critical points are where \(f'(x) = 3x^2 - 3 = 0\), or \(x = -1, x = +1\).
- \(f''(1) = 6 > 0\): local min
- \(f''(-1) = -6 < 0\): local max
- Actual maximum: \(f(3) = 18 > 2 = f(-1)\).

Example

\(f(x) = -|x|\) is not differentiable at its max, \(x = 0\).
1D review

The critical points of \( f(\cdot) \) are the points where \( f'(x) = 0 \).

A critical point is

▶ Local maximum iff''(x) < 0, i.e., \( f''(-1) = -6 < 0 \): local max
▶ Actual maximum: \( f(3) = 18 > 2 = f(-1) \).

Example

\( f(x) = -|x| \) is not differentiable at its max, \( x = 0 \).

---

Lecture 18: Critical points and optimization

Optimization

▷ First derivatives

Second derivatives

Checking first derivative

The critical points of \( f(\cdot) \) are the points where \( f'(x) = 0 \).

A critical point is

▶ Local maximum iff''(x) < 0, i.e., \( f''(-1) = -6 < 0 \): local max
▶ Actual maximum: \( f(3) = 18 > 2 = f(-1) \).

Example

\( f(x) = -|x| \) is not differentiable at its max, \( x = 0 \).

Theorem

If \( f(\cdot, \cdot) \) is diff. at \((a, b)\) and has a local max or min there, then \( \nabla f(a, b) = 0 \).

Proof idea.

If \( \nabla f(a, b) \neq 0 \), then \( \mathcal{D}_{(h, k)} f(a, b) > 0 \) for \((h, k)\) pointing in direction of \( \nabla f \).

Stepping in this direction will increase \( f \):

\[ f(a + \epsilon h, b + \epsilon k) \approx f(a, b) + \epsilon \mathcal{D}_{(h, k)} f(a, b) > f, \]

So, to find local maxes/mins, only need to check where \( \nabla f = 0 \) (assuming \( f \) is differentiable).

Example

Local min of \( f(x, y) = x^2 + y^2 \) occurs where \( \nabla f = (2x, 2y) = 0 \).

Example

\( f(x, y) = |x| + |y| \) has a local min at \((0, 0)\), but is not differentiable there.

Example

For \( f(x, y) = \cos(\pi x) + \cos(\pi y) \), have

\[ \nabla f(x, y) = (-\pi \sin(\pi x), -\pi \sin(\pi y)) \]

Local max/min can only be where \( \nabla f = 0 \), i.e.,

\( \sin(\pi x) = 0 \) (so \( x \) an integer) and \( \sin(\pi y) = 0 \) (so \( y \) an integer).
Critical points

Definition
A critical point of \( f(\cdot, \cdot) \) is a point \((a, b)\) where \( \nabla f(a, b) = 0 \).
This is same as saying \( f_x = f_y = 0 \). Note: two equations in two unknowns.

Exercises
Find the critical points of
- \( f(x, y) = x^4 - x^2 + y^2 + 2xy - 2 \).
- \( f(x, y) = \cos(x) \sin(y) \).

Lecture 18: Critical points and optimization

Optimization

First derivatives

- Second derivatives

Critical points

Definition
A critical point of \( f(\cdot, \cdot) \) is a point \((a, b)\) where \( \nabla f(a, b) = 0 \).
This is same as saying \( f_x = f_y = 0 \). Note: two equations in two unknowns.

Exercises
Find the critical points of
- \( f(x, y) = x^4 - x^2 + y^2 + 2xy - 2 \).
- \( f(x, y) = \cos(x) \sin(y) \).

Answers
- \( \nabla f(x, y) = (4x^3 - 2x + 2y, 2y + 2x) \). This is 0 at \((0, 0)\), \((1, -1)\), and \((-1, 1)\).
- \( \nabla f(x, y) = (-\sin(x) \sin(y), \cos(x) \cos(y)) \).
  For this to be 0, must have either \( \sin x = 0 \) and \( \cos y = 0 \) or \( \sin y = 0 \) and \( \cos x = 0 \).
This happens when \( x = k\pi \) and \( y = (l + 1/2)\pi \) or \( x = (k + 1/2)\pi \) and \( y = l\pi \) for integers \( k, l \).

Quadratic functions

To understand the behaviour near a critical point, need to look at the second derivatives, like in 1D.
Let’s first consider the simplest functions with non-zero second derivatives: quadratic functions, involving only terms \( x^2 \), \( y^2 \), \( xy \), and lower terms.
If \( f(x, y) \) is quadratic, then \( z = f(x, y) \) defines a quadratic surface (Lecture 6). May be
- elliptic paraboloid (e.g., \( z = x^2 + y^2 \)) or
- hyperbolic paraboloid (e.g., \( z = x^2 - y^2 \)), also called saddle surface.

Examples
What type of surface is the graph of the functions below? Is there a local max, local min, or neither?
- \( x^2 + y^2 \)
- \( x^2 - y^2 \)
- \( xy \)
- \( -x^2 - 2y^2 + x \)
- \( -x^2 - 2xy - 2y^2 \)
Classifying quadratic functions

Recall from Lecture 6 (or earlier!):

**Theorem**

The type of a conic section
\[ f(x, y) = Ax^2 + Bxy + Cy^2 = K \]
depends on \( B^2 - 4AC \) (if non-degen.):
- \( B^2 - 4AC < 0 \): ellipse
- \( B^2 - 4AC = 0 \): parabola
- \( B^2 - 4AC > 0 \): hyperbola

Correspondingly, \( z = f(x, y) \) is
- elliptic paraboloid (with max or min) if \( B^2 - 4AC < 0 \)
- hyperbolic paraboloid (with saddle) if \( B^2 - 4AC > 0 \)

**Examples**
- \( x^2 + y^2 \)
- \( x^2 - y^2 \)
- \( xy \)
- \( -x^2 - 2y^2 \)
- \( -x^2 - 2xy - 2y^2 \)

Classifying quadratic functions

Recall from Lecture 6 (or earlier!):

**Theorem**

The type of a conic section
\[ f(x, y) = Ax^2 + Bxy + Cy^2 = K \]
depends on \( B^2 - 4AC \) (if non-degen.):
- \( B^2 - 4AC < 0 \): ellipse
- \( B^2 - 4AC = 0 \): parabola
- \( B^2 - 4AC > 0 \): hyperbola

Correspondingly, \( z = f(x, y) \) is
- elliptic paraboloid (with max or min) if \( B^2 - 4AC < 0 \)
- hyperbolic paraboloid (with saddle) if \( B^2 - 4AC > 0 \)

**Examples**
- \( x^2 + y^2 \): \( B^2 - 4AC = -4 < 0 \), min
- \( x^2 - y^2 \)
- \( xy \)
- \( -x^2 - 2y^2 \)
- \( -x^2 - 2xy - 2y^2 \)
Classifying quadratic functions

Recall from Lecture 6 (or earlier!):

Theorem
The type of a conic section
\[ f(x, y) = Ax^2 + Bxy + Cy^2 = K \]
depends on \( B^2 - 4AC \) (if non-degen.):
- \( B^2 - 4AC < 0 \): ellipse
- \( B^2 - 4AC = 0 \): parabola
- \( B^2 - 4AC > 0 \): hyperbola

Correspondingly, \( z = f(x, y) \) is
- elliptic paraboloid (with max or min) if \( B^2 - 4AC < 0 \)
- hyperbolic paraboloid (with saddle) if \( B^2 - 4AC > 0 \)

Examples
- \( x^2 + y^2 \): \( B^2 - 4AC = -4 < 0 \), min
- \( x^2 - y^2 \): \( B^2 - 4AC = 4 > 0 \), saddle
- \( xy \): \( B^2 - 4AC = 1 > 0 \), saddle
- \(-x^2 - 2y^2 \): \( B^2 - 4AC = -8 < 0 \), max
- \(-x^2 - 2xy - 2y^2 \):

Second derivative test

Example
If \( f(x, y) = x^4 - x^2 + y^2 + 2xy - 2 \), found critical points at \( (1, -1), (0, 0) \), and \((-1, 1)\).
- At \((0, 0)\), \( D = (-2)(2) - 2^2 < 0 \), saddle point.
- At \((1, -1)\), \( D = (10)(2) - 2^2 > 0 \), local min.
- \((-1, 1)\) same as \((1, -1)\).

So this agrees.

Classifying quadratic functions

Recall from Lecture 6 (or earlier!):

Theorem
The type of a conic section
\[ f(x, y) = Ax^2 + Bxy + Cy^2 = K \]
depends on \( B^2 - 4AC \) (if non-degen.):
- \( B^2 - 4AC < 0 \): ellipse
- \( B^2 - 4AC = 0 \): parabola
- \( B^2 - 4AC > 0 \): hyperbola

Correspondingly, \( z = f(x, y) \) is
- elliptic paraboloid (with max or min) if \( B^2 - 4AC < 0 \)
- hyperbolic paraboloid (with saddle) if \( B^2 - 4AC > 0 \)

Examples
- \( x^2 + y^2 \): \( B^2 - 4AC = -4 < 0 \), min
- \( x^2 - y^2 \): \( B^2 - 4AC = 4 > 0 \), saddle
- \( xy \): \( B^2 - 4AC = 1 > 0 \), saddle
- \(-x^2 - 2y^2 \): \( B^2 - 4AC = -8 < 0 \), max
- \(-x^2 - 2xy - 2y^2 \):

Second derivative test

Theorem
If \( f(\cdot, \cdot) \) is twice different'ble and \((a, b)\) is a critical point, let
\[ D = f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2. \]
- If \( D > 0 \) and \( f_{xy}(a,b) > 0 \), then \((a, b)\) is local min.
- If \( D > 0 \) and \( f_{xy}(a,b) < 0 \), then \((a, b)\) is local max.
- If \( D < 0 \), then \((a, b)\) is saddle point.

Example
If \( f(x, y) = x^4 - x^2 + y^2 + 2xy - 2 \), found critical points at \((1, -1), (0, 0)\), and \((-1, 1)\).
- At \((0, 0)\), \( D = (-2)(2) - 2^2 < 0 \), saddle point.
- At \((1, -1)\), \( D = (10)(2) - 2^2 > 0 \), local min.
- \((-1, 1)\) same as \((1, -1)\).

So this agrees.
Second derivative test

Theorem
If \( f(\cdot, \cdot) \) is twice differentiable and \((a, b)\) is a critical point, let
\[
D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2.
\]
- If \( D > 0 \) and \( f_{xx}(a, b) > 0 \), then \((a, b)\) is local min.
- If \( D > 0 \) and \( f_{xx}(a, b) < 0 \), then \((a, b)\) is local max.
- If \( D < 0 \), then \((a, b)\) is saddle point.

\( D \) behaves a little like second derivative.

Note: if \( f \) behaves a little like second derivative, then
\[
D = (−2)(2) − 2^2 < 0.
\]

So this agrees.

Second derivative test

Theorem
If \( f(\cdot, \cdot) \) is twice differentiable and \((a, b)\) is a critical point, let
\[
D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2.
\]
- If \( D > 0 \) and \( f_{xx}(a, b) > 0 \), then \((a, b)\) is local min.
- If \( D > 0 \) and \( f_{xx}(a, b) < 0 \), then \((a, b)\) is local max.
- If \( D < 0 \), then \((a, b)\) is saddle point.

\( D \) behaves a little like second derivative.

Note: if \( f \) behaves a little like second derivative, then
\[
D = (−2)(2) − 2^2 > 0.
\]

So this agrees.

Proof of derivative test

Theorem
\[
D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2.
\]
- If \( D > 0 \), have local min or max.
- If \( D < 0 \), have saddle point.

Proof idea.
Consider slicing in direction of some unit vector \( \vec{u} = (h, k) \).
- If \( D_{\vec{u}}(D_{\vec{u}})(a, b) > 0 \) always, local min.
- If \( D_{\vec{u}}(D_{\vec{u}})(a, b) < 0 \) always, local max.
- If \( D_{\vec{u}}(D_{\vec{u}})(a, b) \) sometimes > 0, sometimes < 0, saddle point.

\[
D_{\vec{u}}(D_{\vec{u}}) = D_{\vec{u}}(hf_x + kf_y)
= hf_xf_{xx} + k(hf_x + kf_y)_y
= h^2f_{xx} + 2hkf_{xy} + k^2f_{yy}.
\]

This can be zero for some \((h, k) \leftrightarrow (f_{xx})(f_{yy}) - (f_{xy})^2 < 0.\]

\[
D_{\vec{u}}(D_{\vec{u}}) = D_{\vec{u}}(hf_x + kf_y)
= hf_xf_{xx} + k(hf_x + kf_y)_y
= h^2f_{xx} + 2hkf_{xy} + k^2f_{yy}.
\]

This can be zero for some \((h, k) \leftrightarrow (f_{xx})(f_{yy}) - (f_{xy})^2 < 0.\]

Proof of derivative test

Theorem
\[
D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2.
\]
- If \( D > 0 \), have local min or max.
- If \( D < 0 \), have saddle point.

Proof idea.
Consider slicing in direction of some unit vector \( \vec{u} = (h, k) \).
- If \( D_{\vec{u}}(D_{\vec{u}})(a, b) > 0 \) always, local min.
- If \( D_{\vec{u}}(D_{\vec{u}})(a, b) < 0 \) always, local max.
- If \( D_{\vec{u}}(D_{\vec{u}})(a, b) \) sometimes > 0, sometimes < 0, saddle point.

\[
D_{\vec{u}}(D_{\vec{u}}) = D_{\vec{u}}(hf_x + kf_y)
= hf_xf_{xx} + k(hf_x + kf_y)_y
= h^2f_{xx} + 2hkf_{xy} + k^2f_{yy}.
\]

This can be zero for some \((h, k) \leftrightarrow (f_{xx})(f_{yy}) - (f_{xy})^2 < 0.\]
Proof of derivative test

Theorem
\[ D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2. \]

- If \( D > 0 \), have local min or max.
- If \( D < 0 \), have saddle point.

Proof idea.
Consider slicing in direction of some unit vector \( \vec{u} = (h, k) \).
- If \( D_\vec{u}(D_{\vec{u}}f)(a, b) > 0 \) always, local min.
- If \( D_\vec{u}(D_{\vec{u}}f)(a, b) < 0 \) always, local max.
- If \( D_\vec{u}(D_{\vec{u}}f)(a, b) \) sometimes > 0, sometimes < 0, saddle point.

\[ D_\vec{u}(D_{\vec{u}}f) = D_\vec{u}(h f_x + k f_y) \]
\[ = h(h f_x + k f_y)_x + k(h f_x + k f_y)_y \]
\[ = h^2 f_{xx} + 2h k f_{xy} + k^2 f_{yy}. \]

This can be zero for some \( (h, k) \) \( \iff \) \( (f_{xx})(f_{yy}) - (f_{xy})^2 < 0 \).

Proof of derivative test

Theorem
\[ D = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2. \]

- If \( D > 0 \), have local min or max.
- If \( D < 0 \), have saddle point.

Proof idea.
Consider slicing in direction of some unit vector \( \vec{u} = (h, k) \).
- If \( D_\vec{u}(D_{\vec{u}}f)(a, b) > 0 \) always, local min.
- If \( D_\vec{u}(D_{\vec{u}}f)(a, b) < 0 \) always, local max.
- If \( D_\vec{u}(D_{\vec{u}}f)(a, b) \) sometimes > 0, sometimes < 0, saddle point.

\[ D_\vec{u}(D_{\vec{u}}f) = D_\vec{u}(h f_x + k f_y) \]
\[ = h(h f_x + k f_y)_x + k(h f_x + k f_y)_y \]
\[ = h^2 f_{xx} + 2h k f_{xy} + k^2 f_{yy}. \]

This can be zero for some \( (h, k) \) \( \iff \) \( (f_{xx})(f_{yy}) - (f_{xy})^2 < 0 \).