Lecture 7   17 july 2013

Chain Rule & Total Deriv

Unconstrained Optimization

Second Derivative Test

Lagrange Multipliers

High dimensional Chain Rule & Total Derivative

Recall $f(x+\Delta x, y+\Delta y) \approx f(x,y) + \partial_x f(x,y)\Delta x + \partial_y f(x,y)\Delta y$

Chain Rule: $\partial_a\left(f(x(a,b), y(a,b))\right) = \partial_x f\, \partial_a x + \partial_y f\, \partial_a y$

$$= \frac{\partial f}{\partial x}\frac{\partial x}{\partial a} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial a}$$

Similarly for $\partial_b f(x(a,b), y(a,b))$.

how much f changes per change in x

how much x changes per change in a

Total derivative. Let $f(t,x,y,z)$ be temperature of ocean at time t & pos x,y,z.

If a submarine travels along curve $x(t), y(t), z(t)$, what is total derivative of temp wrt time?

$$\frac{d}{dt}\left[f(t, x(t), y(t), z(t))\right] = \partial_t f + \underbrace{\partial_x f \frac{dx}{dt} + \partial_y f \frac{dy}{dt} + \partial_z f \frac{dz}{dt}}_{\text{other vars } x,y,z \text{ are not held constant}}$$

Compare to partial deriv $\partial_t f$ – other vars held constant

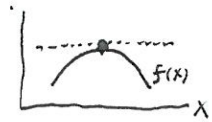
Unconstrained Optimization

If (x,y) is a local min or max of $f(x,y)$

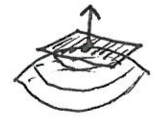
then $\vec{\nabla} f(x,y) = 0$ (assuming diff'ability)

Comments:

- Like 1d: if x is local min $f(x)$, $f'(x)=0$

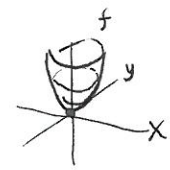


- Tangent plane is flat



- If $\vec{\nabla} f(x,y) = 0$ then (x,y) is critical point of f

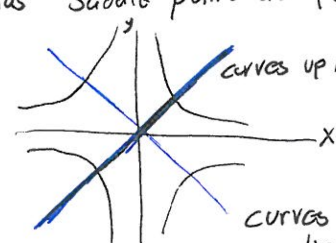
May be max / min / saddle / other.



Examples: $f(x,y) = x^2+y^2$ has local min at $(0,0)$

$f(x,y) = xy$ has saddle point at $(0,0)$

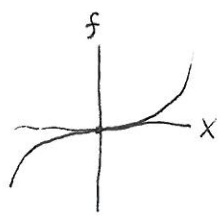
Level sets



curves up in this direction

curves down in this direction

$\cdot\, f(x,y) = x^5$



Not saddle. at $(0,0)$

Second Derivative in 2d

Let (x,y) be critical point of $f(x,y)$. $(\vec{\nabla} f(x,y)=0)$

Hessian matrix H

If $\det\begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{pmatrix} > 0$ & $f_{xx}(x,y) > 0$, local min

If $\det\begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{pmatrix} > 0$ & $f_{xx}(x,y) < 0$, local max

If $\det\begin{pmatrix} \cdots \end{pmatrix} < 0$, saddle point.

Otherwise, inconclusive

Example:

$f(x,y) = x^2+y^2$ at $(0,0)$.

$\nabla f(x,y) = (2x, 2y)$. $\nabla f(0,0) = (0,0)$.

$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. $\det H = 4 > 0$. Local max or min. $f_{xx} > 0$, local min

$f(x,y) = xy$

$\nabla f(x,y) = (y, x)$ $\nabla f(0,0) = (0,0)$.

$H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $\det H = -1 < 0$. Saddle.

Unconstrained Optimization Problems

To find min / max of $f(x,y)$

- Find critical points, where $\vec{\nabla} f(x,y) = 0$
- If necessary, use 2^{nd} deriv test to verify that it is max/min.

Example: $f(x,y) = x^2 - x + y^2 + y$

$\nabla f(x,y) = (2x-1,\ 2y+1) = (0,0) \Rightarrow \begin{array}{l} x = \frac{1}{2} \\ y = -\frac{1}{2} \end{array}$

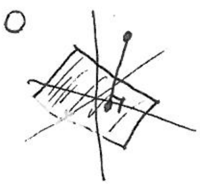
Critical point $(\frac{1}{2}, -\frac{1}{2})$

$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ has pos det. Local min

Posing Unconstrained Optimization

- Quantify Search space
  - If there is a constraint, remove it and lose one degree of freedom
- Quantify objective
  - Simplify (remove/add log or $\sqrt{\ }$ or e^x or $(\cdot)^2$)
- Find critical points

Example: Find the point on the plane $x+y+z=0$ that is closest to $(1,2,3)$.



Need a function measuring distance of an arbitrary point on the plane to $(1,2,3)$

Distance of (x,y,z) to $(1,2,3)$ is $\sqrt{(x-1)^2+(y-2)^2+(z-3)^2}$

The plane is constraint. Given x,y, it specifies z.
Remove z altogether $\quad z=-x-y$

$$\min \sqrt{(x-1)^2+(y-2)^2+(-x-y-3)^2}$$

But why minimize distance. Min distance squared

$$\boxed{\min \quad (x-1)^2+(y-2)^2+(x+y+3)^2}$$

$$f(x,y)=(x-1)^2+(y-2)^2+(x+y+3)^2$$

$$\nabla f = \left(2(x-1)+2(x+y+3),\ 2(y-2)+2(x+y+3)\right)$$

$$= \left(4x+2y+4,\ 2x+4y+2\right)=(0,0) \implies \begin{array}{l} 4x+2y+4=0 \\ 2x+4y+2=0 \end{array}$$

$$\implies \begin{array}{l} x=-1 \\ y=0 \end{array}$$

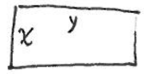
At $x=1, y=0, z=+1$.

Nearest point is $(-1,0,+1)$

Lagrange Multipliers - Constrained optimization

Can solve constrained optimization problems directly

Ex: Find rectangle of largest area given perimeter P.



max xy subject to $2x+2y = P$.

- could solve for y & view as unconstrained problem
- easy in this case, $\underbrace{\text{tough}}_{\text{cumbersome}}$ in some cases (esp if constraint nonlinear)

Standard form:

max/min $f(x,y)$ subject to $g(x,y)=0$

To solve with Lagrange Multipliers:

- Introduce variable λ (lagrange multiplier)
- Form Lagrangian $\mathcal{L}(x,y,\lambda) = f(x,y) + \lambda g(x,y)$.
- Search for critical points in (x,y,λ)
  - Set $\partial_x \mathcal{L} = 0$
    $\partial_y \mathcal{L} = 0$
    $\partial_\lambda \mathcal{L} = 0$.
  - solve for x, y, λ.

Rectangle of Largest area given perim P

Ex: max xy s.t. $2x+2y-P=0$

$$\mathcal{L}(x,y,\lambda) = xy + \lambda(2x+2y-P)$$

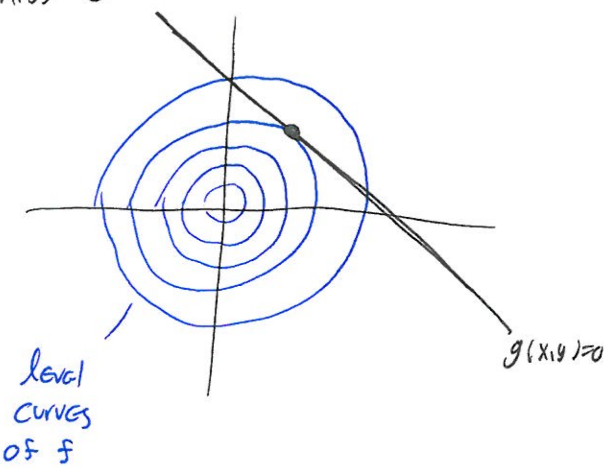
$$\partial_x \mathcal{L} = y + 2\lambda = 0$$
$$\partial_y \mathcal{L} = x + 2\lambda = 0$$
$$\partial_\lambda \mathcal{L} = 2x+2y-P = 0$$

$$\Rightarrow \begin{array}{l} y = -2\lambda \\ x = -2\lambda \end{array} \Rightarrow x = y \Rightarrow \text{Square!}$$

$$x = y = P/4$$

Justification of Lagrange Multipliers

max/min $f(x,y)$ st $g(x,y)=0$



$\mathcal{L} = f(x,y) + \lambda g(x,y)$

$\nabla \mathcal{L} = 0 \Rightarrow \nabla f(x,y) + \lambda \nabla g(x,y) = 0$

∇f is parallel to ∇g at constrained extremum

So level curves of f & g are tangent.

If level curves of f & g weren't tangent, you could move along constraint and increase OR decrease objective