Critical points

Def: $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}$ differentiable, $a \in U$
We say that a is a critical point of f if $\nabla f(a) = \vec{0}$

Def: $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}$, $a \in U$
We say that a is a local min of f if

$$\exists r > 0, \forall x \in U, \|x - a\| < r \Rightarrow f(a) \leq f(x)$$

ie: $\exists r > 0, \forall x \in B(a, r) \cap U, \ f(a) \leq f(x)$

Comment: since $B(a,r) \cap U$ is open as the intersection of two open sets we may assume that $B(a,r) \subset U$ up to shrinking r

Def: We say that a is a local max of f if

$$\exists r > 0, \forall x \in U, \|x - a\| < r \Rightarrow f(a) \geq f(x)$$

ie $\exists r > 0, \forall x \in B(a, r) \cap U, \ f(a) \geq f(x)$

Def: local extremum := local min or local max

Theorem (First derivative test)

Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}$ differentiable, $a \in U$.
If a is a local extremum then it is a critical point.

ie: a local extremum $\Rightarrow$ a critical point: $\left\{\begin{array}{l}\text{the local extrema are} \\ \text{among the critical points}\end{array}\right.$

Δ Let $a = (a_1, \dots, a_n) \in U$ be a local extremum of f
Then a_j is a local extremum of $g(t) = f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n)$

hence $g'(a_j) = 0$ by MAT137, but $g'(a_j) = \frac{\partial f}{\partial x_j}(a)$

Therefore $\nabla f(a) = \vec{0}$ $\square$

Study up to order 2

$U \subset \mathbb{R}^m$ open, $f : U \to \mathbb{R}$ of class C^2, $a \in U$ critical point

then $f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h^t \mathcal{H}_f(a)\, h + E(h)$ where $\frac{E(h)}{\|h\|^2} \xrightarrow[h \to 0]{} 0$

becomes $f(a+h) - f(a) = \frac{1}{2} h^t \mathcal{H}_f(a)\, h + E(h)$ (*)

and by Clairaut's theorem $\mathcal{H}_f(a)$ is a symmetric matrix

Definitions: $A \in M_{m,m}(\mathbb{R})$ symmetric (ie $A^t = A$) is said to be

- positive definite if $\forall h \in M_{m,1}(\mathbb{R})$, $h \neq 0 \Rightarrow h^t A h > 0$
- nonnegative definite if $\forall h \in M_{m,1}(\mathbb{R})$, $h^t A h \geq 0$
- negative definite if $h \neq 0 \Rightarrow h^t A h < 0$
- non positive definite if $\forall h$, $h^t A h \leq 0$
- non-definite if it is not nonnegative definite neither nonpositive definite
  ie $\exists h, k$ s.t. $h^t A h < 0 < k^t A k$
- degenerate if $\det A = 0$
- non-degenerate if $\det A \neq 0$

Theorem: ① positive definite $\Leftrightarrow$ eigenvalues are > 0
② nonnegative definite $\Leftrightarrow$ eigenvalues are ≥ 0
③ negative definite $\Leftrightarrow$ eigenvalues are < 0
④ nonpositive definite $\Leftrightarrow$ eigenvalues are ≤ 0
⑤ indefinite $\Leftrightarrow$ some eigenvalues are > 0 and some are < 0

Δ ① $\Rightarrow$ let λ be an eigenvalue with eigenvector $h \neq 0$ then
$0 < h^t A h = h^t \lambda h = \lambda h^t h = \lambda \|h\|^2 \Rightarrow \lambda > 0$

$\Leftarrow$ recall that, since A is symmetric, we may find an orthogonal basis of $\mathbb{R}^m$ made of eigenvectors of A (**)

□

Corollary: ① positive definite $\Leftrightarrow$ non-degenerate + nonnegative definite
② negative definite $\Leftrightarrow$ non-degenerate + non positive definite

Δ The determinant of a symmetric matrix is the product of its eigenvalues (with mult)

indeed by (**) $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}$ in some basis and then $\det A = \lambda_1 \cdots \lambda_m$ □

Lemma: Let $A \in M_{m,m}(\mathbb{R})$ be a symmetric matrix then

① A positive definite $\Leftrightarrow \exists \lambda > 0, \forall h \in \mathbb{R}^m, h^t A h \geq \lambda \|h\|^2$

② A negative definite $\Leftrightarrow \exists \lambda < 0, \forall h \in \mathbb{R}^m, h^t A h \leq \lambda \|h\|^2$

Δ ① $\Leftarrow$ let $h \in \mathbb{R}^m \setminus \{\vec{0}\}$ then $h^t A h \geq \lambda \|h\|^2 > 0$

$\Rightarrow$ let $(u_1, \dots, u_m)$ be an orthogonal basis of eigenvectors (exists since A is symmetric)
let $h \in \mathbb{R}^m$ then $h = \sum_{i=1}^{m} \alpha_i u_i$, $\alpha_i \in \mathbb{R}$

$$h^t A h = \sum_{i,j} \alpha_i \alpha_j u_i^t A u_j$$

$$= \sum_{i,j} \alpha_i \alpha_j \lambda_j u_i^t u_j \quad \text{where } A u_j = \lambda_j u_j,\ \lambda_j > 0$$

$$= \sum_i \alpha_i^2 \lambda_i u_i^t u_i \quad \text{since } i \neq j \Rightarrow u_i^t u_j = u_i \cdot u_j = 0$$

$$\geq \min(\lambda_i) \sum_i \alpha_i^2 u_i^t u_i$$

$$= \min(\lambda_i) \sum_{i,j} \alpha_i \alpha_j u_i^t u_j \quad \text{since } u_i^t u_j = u_i \cdot u_j = 0$$

$$= \min(\lambda_i) \left(\sum \alpha_i u_i\right) \cdot \left(\sum \alpha_j u_j\right)$$

$$= \min(\lambda_i) \|h\|^2$$

② Apply ① to $-A$

$\square$

ie $\nabla f(a) = \vec{0}$

Theorem: (Second derivative test)

$U \subset \mathbb{R}^n$ open, $f: U \to \mathbb{R}$ of class C^2, $a \in U$ critical point

↳ don't forget this assumption

① If $Hf(a)$ is positive definite then a is a local min

② If $Hf(a)$ is negative definite then a is a local max

③ If $Hf(a)$ is indefinite then a is neither a local max nor a local min

Remark: in all the other cases we can't conclude about the nature of "a" simply from $Hf(a)$

Ex: $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2$ has a local min at $(0,0)$

$g: \mathbb{R}^2 \to \mathbb{R}$, $g(x,y) = x^3$ has no local extremum at $(0,0)$

$Hf(\vec{0})$ and $Hg(\vec{0})$ are both non-negative definite

Proof: ① $f(a+h) - f(a) = \frac{1}{2} h^t Hf(a) h + E(h)$ by ⊛

$$\geq \frac{1}{2}\lambda \|h\|^2 + E(h) \text{ for some } \lambda > 0 \text{ by the lemma}$$

$$= \underbrace{\|h\|^2 \left(\frac{\lambda}{2} + \frac{E(h)}{\|h\|^2} \right)}_{\text{hence } > 0 \text{ for } \|h\| \text{ small enough}}$$

$\frac{\lambda}{2} > 0$, $\frac{E(h)}{\|h\|^2} \xrightarrow{h \to 0} 0$

② apply ① to "$-f$"

③ $\exists h, k$ s.t. $h^t Hf(a) h < 0 < k^t Hf(a) k$

$$f(a+th) - f(a) = t^2 \left(h^t Hf(a) h + \|h\|^2 \frac{E(th)}{\|th\|^2} \right) < 0 \text{ for } t \text{ small enough}$$

$f(a+tk) - f(a) > 0$ for t small enough

So f takes some values $> f(a)$ and $< f(a)$ in any ball centered at a

□

<u>Comment</u>: in the case 3 $g(t) = f(a+th)$ has a local max at 0

and $h(t) = f(a+tk)$ has a local min at 0

$\rightsquigarrow$ looks like a saddle.

⚠ don't forget this assumption

<u>Def</u>: $U \subset \mathbb{R}^n$ open, $f : U \longrightarrow \mathbb{R}$ of class C^2, $a \in U$ $\boxed{\text{critical point}}$

We say that a is a saddle point if $H f_f(a)$ is indefinite

The two-variable case

don't forget this assumption

$\mathcal{U} \subset \mathbb{R}^2$ open, $f : \mathcal{U} \to \mathbb{R}$ of class C^2, $a \in \mathcal{U}$ critical point of f.

Set $\alpha = \dfrac{\partial^2 f}{\partial x^2}(a)$, $\beta = \dfrac{\partial^2 f}{\partial x \partial y}(a) = \dfrac{\partial^2 f}{\partial y \partial x}(a)$, $\gamma = \dfrac{\partial^2 f}{\partial y^2}(a)$

by Clairaut's theorem

then $\mathcal{H}f(a) = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$

Since $\mathcal{H}f(a)$ is a symmetric 2×2-matrix, it has two eigenvalues λ_1, λ_2 (that may be equal) and then its determinant is the product of its eigenvalues, ie: $\lambda_1 \lambda_2 = \det(\mathcal{H}f(a)) = \alpha\gamma - \beta^2$

- $\alpha\gamma - \beta^2 > 0$: then either λ_1, λ_2 are both positive and $\mathcal{H}f(a)$ is positive definite or they are both negative and $\mathcal{H}f(a)$ is negative definite.

  ie either $h \neq 0 \Rightarrow h^t \mathcal{H}f(a) h > 0$ (positive definite)
  or $h \neq 0 \Rightarrow h^t \mathcal{H}f(a) h < 0$ (negative definite)

  But $e_1^t \mathcal{H}f(a) e_1 = (1\ 0)\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \alpha$

  So if $\alpha > 0$ then $\mathcal{H}f(a)$ is positive definite and a is a local min of f
  and if $\alpha < 0$ then $\mathcal{H}f(a)$ is negative definite and a is a local max of f

  Comment: $\alpha = 0$ is impossible since $\mathcal{H}f(a)$ is positive or negative definite
  but you can double check: $\alpha = 0 \Rightarrow \det \mathcal{H}f(a) = -\beta^2 \leq 0$ impossible

- $\alpha\gamma - \beta^2 < 0$: one eigenvalue is > 0 and the other < 0
  then $\mathcal{H}f(a)$ is indefinite and a is a saddle point

- $\alpha\gamma - \beta^2 = 0$: one eigenvalue is 0 so $\mathcal{H}f(a)$ can't be positive/negative definite and since there are only two eigenvalues we can't have a positive and a negative eigenvalue so $\mathcal{H}f(a)$ is not indefinite.

  $\Rightarrow$ we can't conclude

We just proved:

<u>Theorem</u> (Monge)

ie $\frac{\partial f}{\partial x}(a) = \frac{\partial f}{\partial y}(a) = 0$

$U \subset \mathbb{R}^2$ open, $f : U \to \mathbb{R}$ of class C^2, $a \in U$ critical point of f

We denote $Hf(a) = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$, $\alpha = \partial_x^2 f(a)$, $\beta = \partial_{xy}^2 f(a) = \partial_{yx}^2 f(a)$, $\gamma = \partial_y^2 f(a)$

- If $\begin{cases} \alpha\gamma - \beta^2 > 0 \\ \alpha > 0 \end{cases}$ then a is a local min of f
- If $\begin{cases} \alpha\gamma - \beta^2 > 0 \\ \alpha < 0 \end{cases}$ then a is a local max of f
- If $\alpha\gamma - \beta^2 < 0$ then a is a saddle point of f
- If $\alpha\gamma - \beta^2 = 0$ then we can't determine the nature of a from $Hf(a)$

<u>Decision tree</u>: