Local Extrema for $f(x_1, x_2.., x_n)$

Set of n variables is represented as a vector

Definition

[Local minimum]: *A function* $f : \mathcal{D} \rightarrow \Re$ *of* n *variables has a local minimum at* $\mathbf{x}^0$ *if* $\exists \mathcal{N}(\mathbf{x}^0)$ *such that* $\forall\ \mathbf{x} \in \mathcal{N}(\mathbf{x}^0)$, $f(\mathbf{x}^0) \leq f(\mathbf{x})$. *In other words,* $f(\mathbf{x}^0) \leq f(\mathbf{x})$ *whenever* $\mathbf{x}$ *lies in some neighborhood around* $\mathbf{x}^0$. *An example neighborhood is the circular disc when* $\mathcal{D} = \Re^n$.

Circular disc: {x s.t ||x-x^0||^2 <= \eps}

Definition

[Local maximum]: .... $f(\mathbf{x}^0) \geq f(\mathbf{x})$.

Local Extrema

These definitions are exactly analogous to the definitions for a function of single variable. Figure 1 shows the plot of $f(x_1, x_2) = 3x_1^2 - x_1^3 - 2x_2^2 + x_2^4$. As can be seen in the plot, the function has several local maxima and minima.

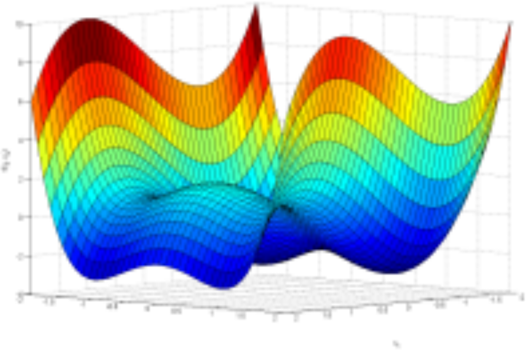


Figure 1:

Convexity and Extremum: Slopeless interpretation (SI)

Definition

A function f is convex on $\mathcal{D}$, *iff*

Convex combination of vectors x1 and x2

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) \tag{1}$$

and is strictly convex on $\mathcal{D}$, *iff*

In 1-d case, we saw convex combination corresponded to points on line segment connecting x1 and x2.

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) \tag{2}$$

whenever $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, $\mathbf{x}_1 \neq \mathbf{x}_2$ and $0 < \alpha < 1$.

Note: This implicitly assumes that whenever $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, then the convex combination is also in D

Convexity and Extremum: Slopeless interpretation (SI)

Definition

A function f is convex on $\mathcal{D}$, *iff*

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \tag{1}$$

and is strictly convex on $\mathcal{D}$, *iff*

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \tag{2}$$

whenever $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, $\mathbf{x}_1 \neq \mathbf{x}_2$ and $0 < \alpha < 1$. Require set D to be convex

Note: This implicitly assumes that whenever $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, $\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \in \mathcal{D}$

Local Extrema

Figure 2 shows the plot of $f(x_1, x_2) = 3x_1^2 + 3x_2^2 - 9$. As can be seen in the plot, the function is cup shaped and appears to be convex everywhere in $\Re^2$. (where the set R^2 itself is convex)

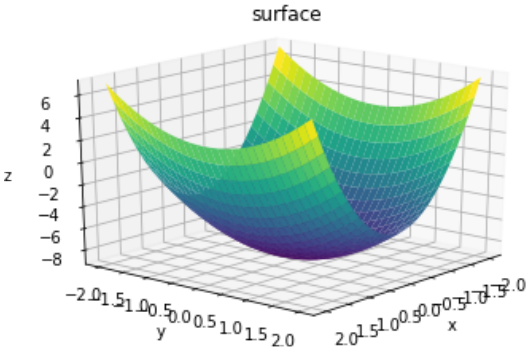


Figure 2:

From $f(x) : \Re \to \Re$ to $f(x_1, x_2 \ldots x_n) : \mathcal{D} \to \Re$

Need to also extend

- Extreme Value Theorem
- Rolle's theorem, Mean Value Theorem, Taylor Expansion
- Necessary and Sufficient first and second order conditions for local/extrema
- First and second order conditions for Convexity

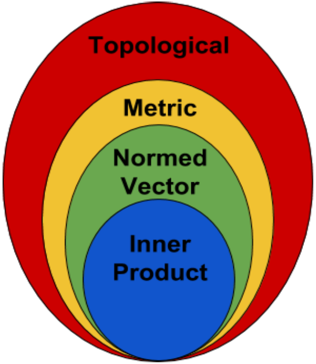
Need following notions/definitions in $\mathcal{D}$

- Neighborhood and open sets/balls ($\Leftarrow$ Local extremum)
- Bounded, Closed Sets ($\Leftarrow$ Extreme value theorem)
- Convex Sets ($\Leftarrow$ Convex functions of n variables)
- Directional Derivatives and Gradients ($\Leftarrow$ Taylor Expansion, all first order conditions)

Spaces and Mathematical Structures

Contents: The Mathematical Structures called Spaces

- Topological Spaces: Notion of neighbourhood of points.
- Metric Spaces: Notion of positive distance between two points.
- Normed Vector Spaces: Notion of positive length of each point.
- Inner Product Spaces: Notion of projection of one point on another, both positive and negative.



Topological Spaces

Set of points ($\mathcal{X}$) along with the set of neighbors ($\mathcal{N}(\mathbf{x})$) of each point ($\mathbf{x}$), with certain axioms required to be satisfied by the points and their neighbors.

- Example1: A topological space is an ordered pair ($\mathcal{X}$,$\mathcal{N}$), where:
  - $\mathcal{X}$ is a set
  - $\mathcal{N}$ is a collection of subsets of $\mathcal{X}$, satisfying the following axioms:
    - The empty set and $\mathcal{X}$ itself belong to $\mathcal{N}$.
    - Any (finite or infinite) union of members of $\mathcal{N}$ still belongs to $\mathcal{N}$.
    - The intersection of any finite number of members of $\mathcal{N}$ still belongs to $\mathcal{N}$.
- As per above example, which out of following are toplogies with $\mathcal{X} = \{1,2,3\}$ and $\mathcal{N} =$
  - $\{\{\},\{1,2,3\}\}$
  - $\{\{\},\{1\},\{1,2,3\}\}$
  - $\{\{\},\{1\},\{2\},\{1,2\},\{1,2,3\}\}$
  - $\{\{\},\{2\},\{1,2\},\{2,3\},\{1,2,3\}\}$
  - $\{\{\},\{1\},\{2\},\{1,2,3\}\}$
  - $\{\{\},\{1,2\},\{2,3\},\{1,2,3\}\}$

Topological Spaces

Set of points ($\mathcal{X}$) along with the set of neighbors ($\mathcal{N}(\mathbf{x})$) of each point ($\mathbf{x}$), with certain axioms required to be satisfied by the points and their neighbors.

- Example1: A topological space is an ordered pair ($\mathcal{X}$,$\mathcal{N}$), where:
  - $\mathcal{X}$ is a set
  - $\mathcal{N}$ is a collection of subsets of $\mathcal{X}$, satisfying the following axioms:
    - ★ The empty set and $\mathcal{X}$ itself belong to $\mathcal{N}$.
    - ★ Any (finite or infinite) union of members of $\mathcal{N}$ still belongs to $\mathcal{N}$.
    - ★ The intersection of any finite number of members of $\mathcal{N}$ still belongs to $\mathcal{N}$.

- As per above example, which out of following are toplogies with $\mathcal{X} = \{1,2,3\}$ and $\mathcal{N} =$
  - $\{\{\},\{1,2,3\}\}$ **Yes**
  - $\{\{\},\{1\},\{1,2,3\}\}$ **Yes**
  - $\{\{\},\{1\},\{2\},\{1,2\},\{1,2,3\}\}$ **Yes**
  - $\{\{\},\{2\},\{1,2\},\{2,3\},\{1,2,3\}\}$ **Yes**
  - $\{\{\},\{1\},\{2\},\{1,2,3\}\}$ **No as $\{1\} \cup \{2\} \notin \mathcal{N}$**
  - $\{\{\},\{1,2\},\{2,3\},\{1,2,3\}\}$ **No as $\{1,2\} \cap \{2,3\} \notin \mathcal{N}$**

Metric Spaces

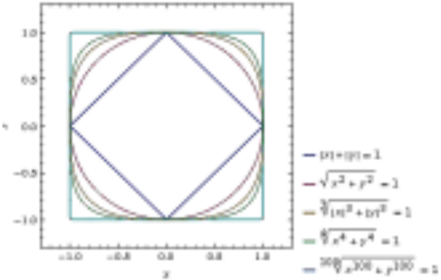
Set of points ($\mathcal{X}$) along with a notion of distance d($\mathbf{x}_1, \mathbf{x}_2$) between any two points($\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}$) such that:

1. d($\mathbf{x}_1, \mathbf{x}_2$) ≥ 0 (non-negativity).
2. d($\mathbf{x}_1, \mathbf{x}_2$) $= 0$ iff $\mathbf{x}_1 = \mathbf{x}_2$ (identity).
3. d($\mathbf{x}_1, \mathbf{x}_2$) $=$ d($\mathbf{x}_2, \mathbf{x}_1$) (symmetry).
4. d($\mathbf{x}_1, \mathbf{x}_2$) $+$ d($\mathbf{x}_2, \mathbf{x}_3$) $\geq$ d($\mathbf{x}_1, \mathbf{x}_3$) (triangle inequality).

Metric Spaces

Examples:

- 1-metric d_1: The plane with the taxi cab metric
  - $d((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)) = |\mathbf{x}1 - \mathbf{x}_2| + |\mathbf{y}_1 - \mathbf{y}_2|$
- 2-metric d_2: The plane R2 with the "usual distance" (measured using Pythagoras's theorem):
  - $d((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)) = \sqrt{((\mathbf{x}_1 - \mathbf{x}_2)^2 + (\mathbf{y}_1 - \mathbf{y}_2)^2)}$.
- Infinity metric d_∞: The plane with the maximum metric
  - $d((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)) = max(|\mathbf{x}_1 - \mathbf{x}_2|, |\mathbf{y}_1 - \mathbf{y}_2|)$.



Normed Vector Spaces

- **Vector Space:** A space consisting of vectors, together with the
  1. **associative and commutative operation of addition of vectors**,
  2. **associative and distributive operation of multiplication of vectors by scalars.**
- **Norm:** A function that assigns a strictly positive length or size to each vector in a vector space — save for the zero vector, which is assigned a length of zero.
- **Normed Vector Space:** A vector space on which a norm is defined.

Normed Vector Spaces

A vector space on which a norm is defined.

- In any real vector space R^n, the length of a vector has the following properties:
  1. The zero vector, 0, has zero length; every other vector has a positive length.
     - ★ $||x|| \geq 0$, and $||x|| = 0$ iff $x = 0$.
  2. Multiplying a vector by a positive number changes its length without changing its direction. Moreover,
     - ★ $||\alpha x|| = |\alpha| \; ||x||$ for any scaler α.
  3. The triangle inequality holds. That is, taking norms as distances, the distance from point A through B to C is never shorter than going directly from A to C, or the shortest distance between any two points is a straight line.
     - ★ $||x_1 + x_2|| \leq ||x_1|| + ||x_2||$ for any vectors x_1 and x_2.

The generalization of these three properties to more abstract vector spaces leads to the notion of norm. For example: A matrix norm.

Additionally, in the case of square matrices (thus, m = n), some (but not all) matrix norms satisfy the following condition, which is related to the fact that matrices are more than just vectors: $||\mathbf{AB}|| \leq ||\mathbf{A}|| \; ||\mathbf{B}||$ **for all matrices A and B in** K^{nxn}**.**