ETMAG

LECTURE 11

- Mean Value Theorem
- Extremal and Inflection Points
- Vector spaces

Theorem (Rolle's Theorem)

If f is continuous on a closed interval $[a, b]$ and differentiable at every point of the open interval (a, b) and $f(a) = f\ (b)$ then there is at least one point $x_0 \in (a, b)$ such that $f'(x_0) = 0$.

Proof.

If f is constant – trivial. If not, $f(a) \neq \inf(f([a, b])$ or $f(a) \neq \sup(f([a, b])$. Suppose (without losing generality) that $f(a) \neq \inf(f([a, b])$. Since a function continuous on a closed interval assumes its extreme values, there exists $x_0 \in [a, b]$ such that $f(x_0) = \inf(f([a, b])$ and $x_0 \neq a, x_0 \neq b$. Hence f has a local minimum at x_0 and $f'(x_0) = 0$. QED

Theorem (Mean Value Theorem)

If f is continuous on a closed interval $[a, b]$ and differentiable at every point of (a, b) then there is at least one point $x_0 \in (a, b)$ such that $f'(x_0) = \frac{f(b)-f(a)}{b-a}$.

Proof.

Let $h(x) = \frac{f(b)-f(a)}{b-a}(x - a) + f(a)$, i.e., $y = h(x)$ is the line through $\big(a, f(a)\big)$ and $\big(b, f(b)\big)$. Consider $g(x) = f(x) - h(x)$.Then g is continuous on $[a, b]$ and is differentiable on (a, b).

$g(a) = f(a) - \frac{f(b)-f(a)}{b-a} 0 - f(a) = 0$ and

$g(b) = f(b) - \frac{f(b)-f(a)}{b-a}(b - a) - f(a) = f(b) - f(b) + f(a) - f(a) = 0$. By Rolle's Theorem there exists x_0 such that $0 = g'(x_0) = f'(x_0) - \frac{f(b)-f(a)}{b-a}$. QED

Corollary 1. (of the mean value theorem)
If $f'(x) = 0$ on an open interval (a, b) then f is constant **on the interval**.

Proof. (by contradiction)

Suppose f is not constant on (a, b). Then, for some $c, d \in (a, b)$ $f(c) \neq f(d)$. From the MVT applied to f on (c, d), there exists $x_0 \in (c, d)$ such that $f'(x_0) = \frac{f(d)-f(c)}{d-c} \neq 0$ because $f(c) \neq f(d)$. QED

Corollary 2.
If $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b) and $f'(x) > 0$ for every $x \in (a, b)$ then $f(x)$ is increasing **on the interval**.

Corollary 3.
If $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b) and $f'(x) < 0$ for every $x \in (a, b)$ then $f(x)$ is decreasing **on the interval**.

Comprehension. Prove Corollaries 2 and 3.

Corollary 4.

Let $f\ (x)$ be differentiable in an open interval containing p and let $f'(p) = 0$. Then

- if $f'(x) > 0$ for $x\ < p$ and $f'(x) < 0$ for $x > p$ then f has a local maximum at p,
- if $f'(x) < 0$ for $x\ < p$ and $f'(x) > 0$ for $x > p$ then f has a local minimum at p,
- if $f'(x)$ doesn't change its sign around p then there f has no extreme value at p.

Proof.

It is an immediate consequence of corollaries 2 and 3. QED

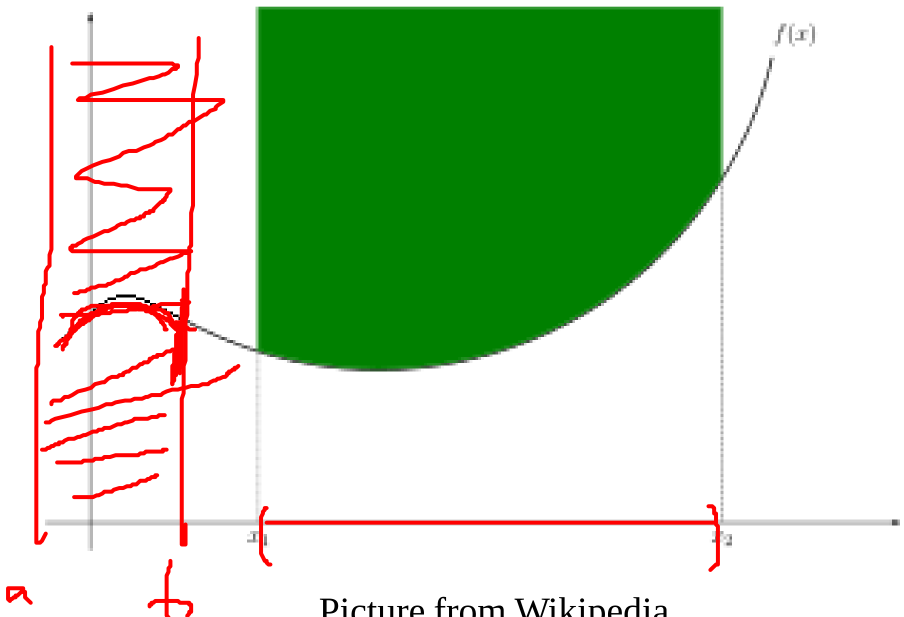
Example.
Discuss monotonicity of $f(x) = x^3 - 3x^2 - 9x + 5$.

$f'(x) = 3x^2 - 6x - 9 = 3(x^2 - 2x - 3) = 3\,(x+1)(x-3)$.
Hence, $f'(x) = 0$ at $x = -1$ and at $x = 3$.
$3\,(x+1)(x-3) < 0$ iff $x \in (-1,3)$ hence, f is decreasing on $(-1,3)$ and increasing on $(-\infty,-1) \cup (3,\infty)$. As a bonus we obtain that f has a locally maximal value of 10 at $x = -1$ and a locally minimal value of -22 at $x = 3$.

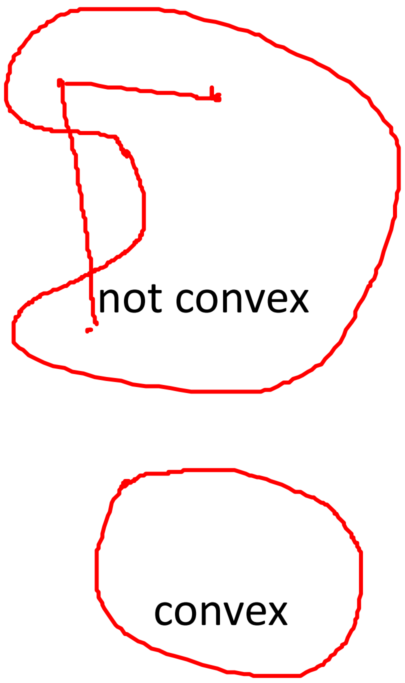
Remark. (on approximation)
The Mean Value Theorem says that (under some conditions), given two points x_0 and x there exists a point $c \in (x_0, x)$ such that $f'(c) = \frac{f(x)-f(x_0)}{x-x_0}$ which means $f(x) - f(x_0) = f'(c)(x - x_0)$ or $f(x) = f(x_0) + f'(c)(x - x_0)$

If we replace $f'(c)$ with $f'(x_0)$ we get an *approximation* of $f(x)$. This strategy can be refined so that one can get better and better approximation of $f(x)$ using derivatives of higher and higher order. Eventually it leads to the Taylor series for a function which you will study in the second semester.

Definition. (convexity of a set)
A set of points S is said to be *convex* iff for every two points $a,b \in S$ the segment joining *a* and *b* is contained in S.



Picture from Wikipedia



Definition. (convexity of a function)
A continuous function is said to be *convex* (*concave* up) on an interval (p, q) iff the area **above** the graph of the function is a convex set.

A continuous function is said to be *concave* (*concave down*) on an interval (p, q) iff the area **below** the graph of the function is a convex set.

Fact.

If $f'(x)$ is increasing on (p, q) then f is convex on (p, q).

If $f'(x)$ is decreasing on (p, q) then f is concave on (p, q).

Proof. (common sense rather than formal)

If $f'(x)$ is increasing on (p, q) then f grows faster and faster which means its graph is bent upward.

Definition.

If convexity of f changes at a point t then t is called *a point of inflection* for f.

The last fact clearly implies the following:

Fact.

If f is twice differentiable on (p, q) then:

- if $f''(x) > 0$ on (p, q) then f is convex on (p, q).
- if $f''(x) < 0$ on (p, q) then f is concave on (p, q).
- if t is a point of inflection for f then $f''(t) = 0$ (but not the other way around).

VECTOR SPACES

Fields

Definition.
An algebra $(X, \#, \&)$ with two binary operations is called a *field* if and only if

1. both operations are commutative,
2. both are associative,
3. there exist identity elements $e_{\#}$ and $e_{\&}$for # and &, resp.
4. every element of X is invertible w.r.t. #
5. every element of X except $e_{\#}$ is invertible w.r.t. &
6. & is distributive over #
7. $|X| \geq 2$ (this can be replaced with $e_{\#} \neq e_{\&}$).

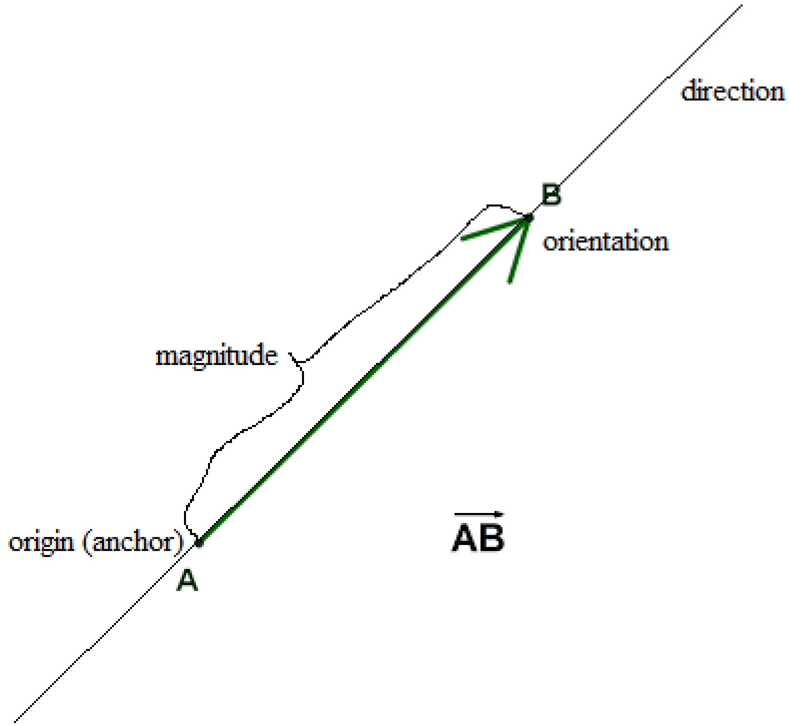
Comments and examples.

- This definition is modeled on the algebra $(\mathbb{R}, +, \cdot)$
- Other examples of fields are $(\mathbb{Q}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$
- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{R}, \cdot, +)$ are not fields (for different reasons).

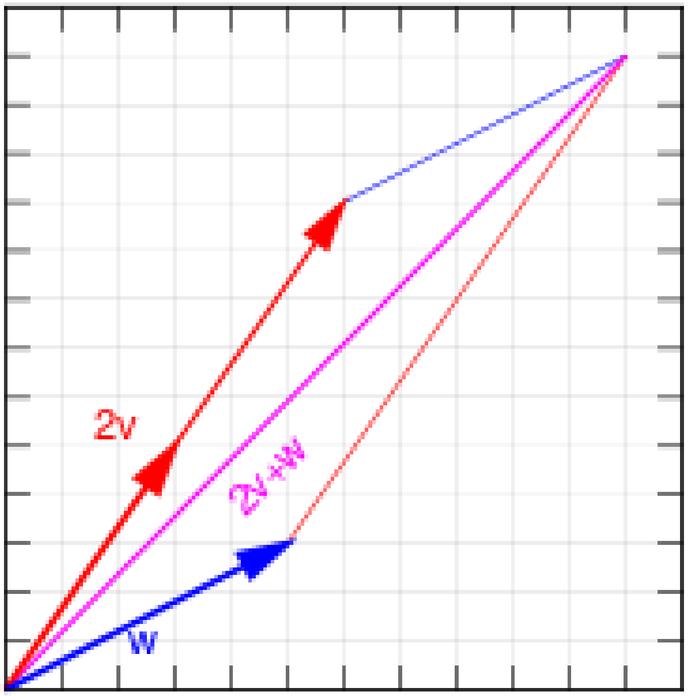
The story of vectors

Vectors often appear in physics where they are used to represent quantities such as a force, the velocity or the acceleration of an object and others, that are not fully representable by a single number like, for example, the mass of an object or the volume of a solid or the area of a plane region. The fact that they are characterized by such properties as the magnitude, the direction, orientation and, often, a point of origin (as in the case of a force) suggests that they may be represented as arrows whose length is proportional to the magnitude. The other attributes like direction, orientation and the anchor point are more or less self-explanatory.

Geometrically, we identify a *vector* with an *ordered pair of points* AB, point A being the *anchor point* or the *origin* of the vector while the location of B depends on the remaining attributes of the quantity which is being represented by the vector. Usually, we place an arrow above AB, $\overrightarrow{AB}$, to denote the vector with the origin A and the endpoint B.



Two vectors anchored at a point *A* can be added using the parallelogram rule. The sum is also anchored at *A*. A vector can be *scaled* by a number, a *scalar*. Scaling preserves the origin and the direction of the vector. It may affect the orientation (if the scalar is negative) and the length (if the scalar is neither 1 nor –1). Hence, in order to create the algebra of vectors we consider the set of vectors anchored at a single point.



In order to use algebraic approach to vectors we consider the space $\mathbb{R}^2$ or $\mathbb{R}^3$ or some such and we assume that all vectors originate at (0, … ,0). Thus, every vector is uniquely identified by a single point namely, its endpoint. This strategy results in a very easy algebraic definition of vector operations. If you have vectors v_1 and v_2 represented by their respective endpoints (a, b) and (c, d) then $v_1 + v_2$ is represented by $(a + c, b + d)$ and $p(a, b)$ by (pa, pb)

We often write $(a, b) + (c, d) = (a + c, b + d)$ and $p(a, b) = (pa, pb)$ but you should be aware that this does not mean that we add or scale points of the plane (or other Euclidean space). We add and scale vectors who by default originate at (0,0) and terminate at (a, b) and (c, d), respectively.

Definition. (Vector space, formal definition)

A *vector space* over a field $\mathbb{K}$ (also called a *linear space*) is an ordered triple $(V, \mathbb{K}, f)$, where

- V is an Abelian group with the operation usually denoted by $+$, whose elements are called *vectors*,
- $\mathbb{K}$ is a field with operations denoted, somewhat confusingly by $+$ and by $\cdot$. Elements of $\mathbb{K}$ are called *scalars*,
- f is a function from $\mathbb{K} \times V$ into V called *scaling*. $f(p, v)$ is often, confusingly, denoted by $p \cdot v$,

and

1. $(\forall \lambda \in \mathbb{K})(\forall u, v \in V)\ \lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$
2. $(\forall \alpha, \beta \in \mathbb{K})(\forall v \in V)\ (\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$
3. $(\forall \alpha, \beta \in \mathbb{K})(\forall v \in V)\ (\alpha \cdot \beta) \cdot v = \alpha \cdot (\beta \cdot v)$
4. $(\forall v \in V)\ 1 \cdot v = v$, where 1 denotes the identity element of the field multiplication (second operation).

Notice the ambiguity caused by the double meaning of the + symbol. This is a BAD, UGLY monster but it is traditional. We let the context decide which "+" means scalar-to-scalar, and which vector-to-vector addition. Otherwise, we would have to introduce extra symbols for scaling and vector addition that could be even more confusing. And it would be harder to type.

Similar remark applies to the dot · , which denotes both scaling (i.e., scalar-by-vector multiplication) and scalar-by-scalar multiplication within the field.

Another problem is caused by the identity elements. People often use 0 to denote the identity element of both the scalar-to-scalar addition and vector-to-vector addition and let the context decide which is which. Other people distinguish between the two using symbols like $\mathbf{0}$ (boldface zero), $\mathbb{0}$ (blackboard bold zero), θ or Θ to denote the "zero vector". That's because sometimes context is not enough, e.g., $0 \cdot 0$ makes sense both in case when the second 0 is the zero scalar and when the second 0 is the zero vector.

Examples.

Let $\mathbb{K}$ be any field and let $n \in \mathbb{N}$ be a natural number. Then

$\mathbb{K}^n = \{(x_1, x_2, \dots, x_n)\colon\ (\forall i = 1,2, \dots n)\ x_i \in \mathbb{K}\}$

together with vector addition defined as follows:

$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$
(*called component-wise addition*)

and scaling

$\alpha(x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$ (*component-wise scaling*)

forms a vector space over $\mathbb{K}$.

The zero vector is $\Theta = (0,0, \dots, 0)$, the inverse of a vector $v = (a_1, a_2, \dots a_n)$ is $(-a_1, -a_2, \dots, -a_n)$.

Notice that in case of $n = 1$ we say (more or less) that every field is a vector space over itself.

Examples.

Let $\mathbb{K}$ be a field. The set $\mathbb{K}[x]$ of all polynomials over $\mathbb{K}$ with the standard polynomial addition and multiplication by a constant from $\mathbb{K}$ forms a vector space over $\mathbb{K}$. The zero vector 0 is the zero polynomial $\mathbf{0}$. Similarly, the set $\mathbb{K}_n[x]$ of polynomials of degree less than or equal to n over the field $\mathbb{K}$ is a vector space.

Example (generalized version of the previous one)

Let $\mathbb{K}$ be a field and let X be a set (any set). Let $V = \mathbb{K}^X$. We define function addition and scaling as the usual operations on functions (i.e., $(f + g)(x) = f(x) + g(x)$, where the second plus denotes operation one in $\mathbb{K}$ and $(\alpha \cdot f)(x) = \alpha \cdot f(x)$). $\mathbb{K}^X$ is a vector space over $\mathbb{K}$.

Proof. V is obviously closed under + (the sum of two functions from V exists and is a function from V). Is $f + (g + h) = (f + g) + h$? WTH does it mean that two functions are equal? They have the same domain and range, which is obvious, and $(\forall x \in \mathrm{X})$ $[f + (g + h)](x) = [(f + g) + h](x)$. The LHS = $f(x) + (g + h)(x) = f(x) + \big(g(x) + h(x)\big) = \big(f(x) + g(x)\big) + h(x) = (f + g)(x) + h(x) = [(f + g) + h](x)$ = RHS.

In the same way we can show that + is commutative. What, if anything, is the zero vector, Θ? We define $\Theta(x) = 0$ for every $x \in X$, where "0" denotes the zero scalar.

Proof. (continued)

The inverse for f (w.r.t. +) is $(-f)$ defined as $(-f)(x) = -(f(x))$, where the second minus denotes the inverse in $\mathbb{K}$ of an element of $\mathbb{K}$ w.r.t. +.

Remaining axioms can be verified in the similar fashion. In each case the identity to be verified boils down to an axiom of fields. E.g.,

scaling

$$(\alpha + \beta) \cdot f = \alpha \cdot f + \beta \cdot f$$

scalar addition   vector addition

follows from

field multiplication

$$(\alpha + \beta) \cdot f(x) = \alpha \cdot f(x) + \beta \cdot f(x) \quad \text{for all } x \in X.$$

field addition

(distributivity law in the field)

Example. (A REALLY outlandish one)

Let X be any set. We will use $V = (2^X, \div)$ as the Abelian group of vectors, where $\div$ denotes the operation of symmetric difference of sets, $A \div B = (A \cup B) \setminus (A \cap B)$. We will also use $(\mathbb{Z}_2, \oplus, \otimes)$ as the field of scalars. Scaling is defined as follows:

for every set A, $0 \cdot A = \varnothing$ and $1 \cdot A = A$.

Comprehension.
Check that $(V, \mathbb{Z}_2, \cdot)$ is a vector space.