# Linear Mathematics.

**Vector spaces, linear methods and tensors throughout STEM.**

**Volume 1 of**

## Widely-Applicable Mathematics Series I.

**Improving everything with a pinch of Combinatorics**

**Edward Anderson***

* Dr.E.Anderson.Maths.Physics *at* protonmail.com

Institute for the Theory of STEM

# Chapter 11

# Convex subsets and functions

## 1.A. Convex subsets of vector spaces

### 11.1 Introduction to Convex subsets

View $\mathbb{R}^n$ once more as a vector space. So it has a well-defined $\underline{\mathbf{0}}$ and thus well-defined position vectors with respect to $\underline{\mathbf{0}}$. And well-defined addition and scalar multiplication operations, as combine to form

$$\lambda\, \boldsymbol{c}_0 \;+\; (1 \;-\; \lambda)\, \boldsymbol{c}_1 \;. \tag{11.1}$$

Next consider a region $\mathfrak{C} \subseteq \mathbb{R}^n$. Let the points $C_0 \neq C_1$ with position vectors $\boldsymbol{c}_0, \boldsymbol{c}_1$ lie within $\mathfrak{C}$. The closed line segment through these points is

$$\Sigma_{0\,1} \;:=\; \{\; \underline{\boldsymbol{c}}_\lambda \;:=\; \lambda\, \underline{\boldsymbol{c}}_0 \;+\; (1 \;-\; \lambda)\, \underline{\boldsymbol{c}}_1 \mid \lambda \;\in\; [\,0,\,1\,] \;\} \;.$$

Or, as a set of points,

$$\{\; C_\lambda \mid \lambda \;\in\; [\,0,\,1\,] \;\} \;.$$

**Definition 1-c** Then $\mathfrak{C}$ is *c-convex* if for all pairs of points $C_0 \neq C_1 \in \mathfrak{A}$, $\Sigma_{0\,1}$ itself $\subseteq \mathfrak{A}$.

**Definition 1-o** A region $\mathfrak{C} \subseteq \mathbb{R}^d$ is *o-convex* if the open line segment, i.e. $\lambda \in (\,0,\,1\,)$ version of the above holds.

**Remark 1** $\mathfrak{C}$ is a subset of $\mathbb{R}^d$, rather than necessarily being a vector subspace. In the present context, $\mathbb{R}^d$ can furthermore be viewed as an Affine space $\mathfrak{A}$.

**Exercise 1** Show that Definitions 1-c and 1-o are equivalent. By which we shall henceforth refer to either as just *Convex* subsets.

♠ ◇ ♣

**Exercise 2** Show that a region $\mathfrak{R}$ being Convex bears no relation to whether $\mathbf{0} \in \mathfrak{R}$.

**Remark 3** We have thus rendered redundant all but 3 of Fig 1.1's subfigures (the boxed ones).

**Exercise 3** Show that a region of $\mathbb{R}^d$ being Convex is unaffected by scalings, rotations or reflections.

♠ ◇ ♣

**Examples 1 to 13** of Convex subsets, and **Examples A to H** of non-convex subsets, all $\subseteq \mathbb{R}^2$, are provided in Fig 1.2.

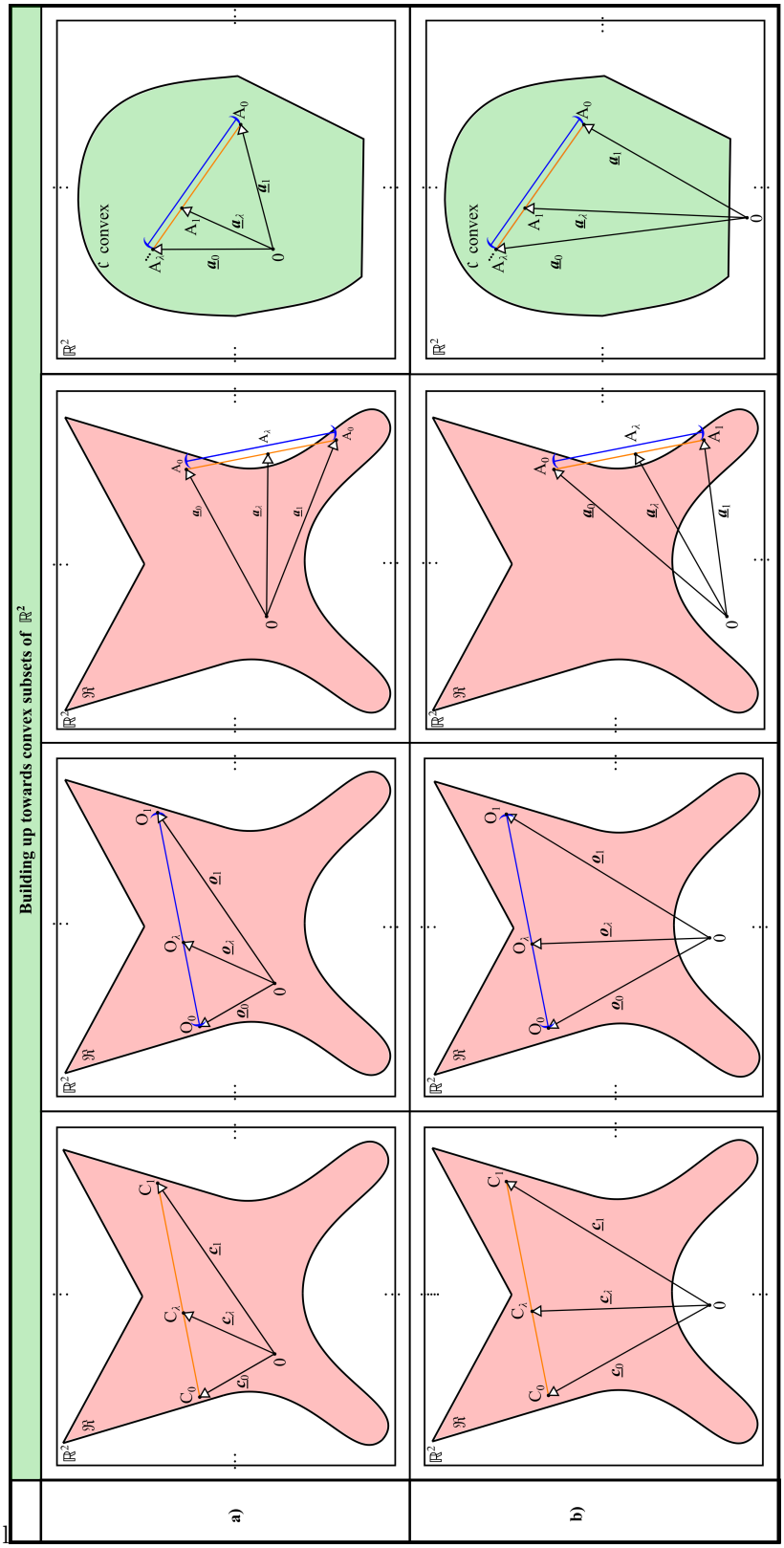


Figure 11.1:

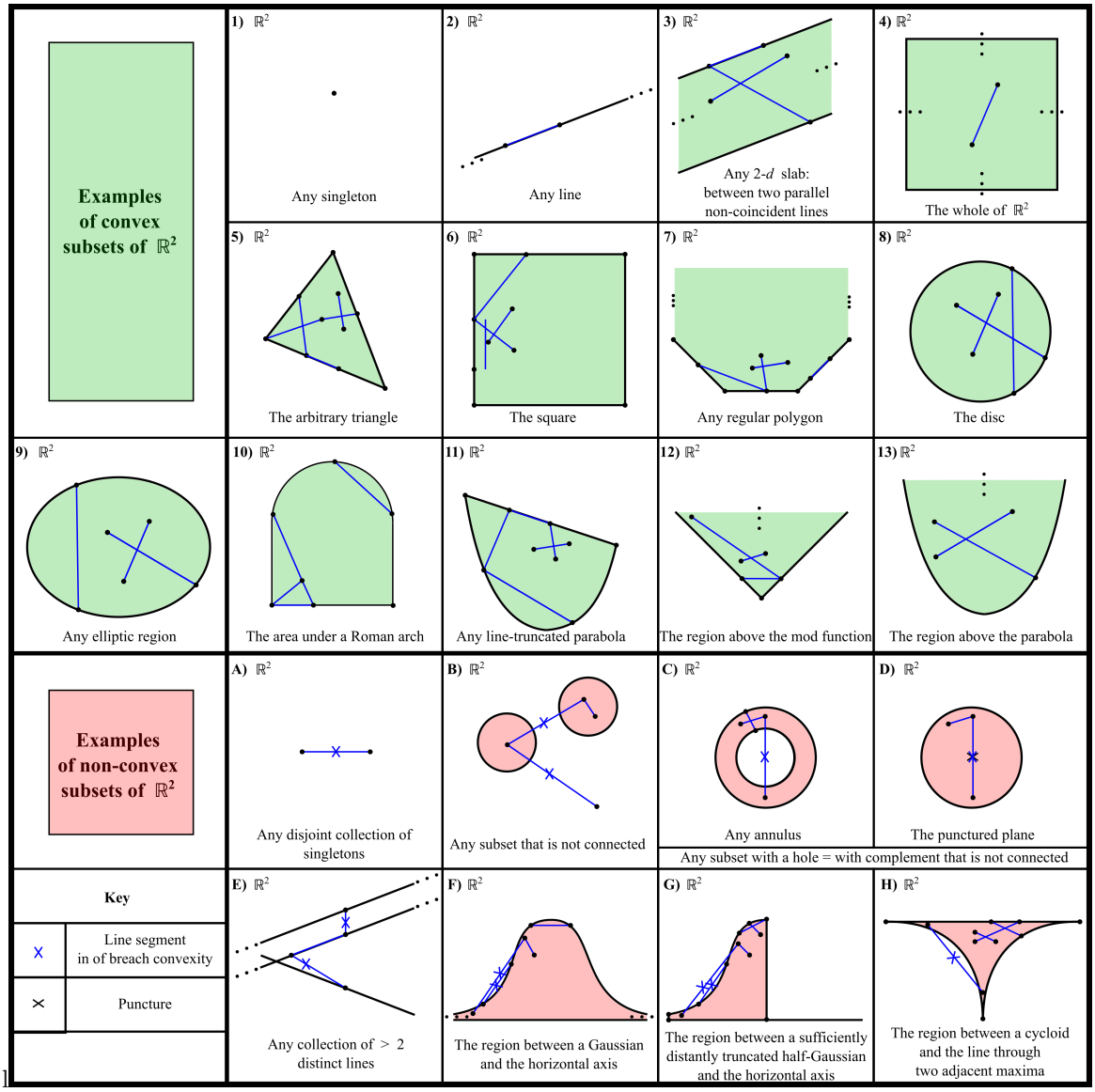


Figure 11.2:

**Exercise 4** Which qualitative types of quadrilateral are Convex, and which are not?

**Remark 3** Convex subset notions extend to any vector space $\mathfrak{V}$ or Affine space $\mathfrak{a}$ (so long as $\mathbb{R}$-LCs remain sufficiently meaningful).

**Exercise 5** (Long) a) Show that Convexity of regions is

i) neither unitive nor disjunctive.

ii) Finitely intersective.

iii) Not hereditary.

iv) Finitely productive.

v) Not quotientive.

b) Also establish which of these adjectives apply to non-convexity of regions.

**Definition 2** Let us also introduce the *arena of Convex subsets* of $\mathfrak{A}$ ,
$$\mathfrak{CS}\text{ubset}\,(\mathfrak{A})\ .$$

**Remark 4** It may be useful to think of $\mathfrak{AS}\text{ubset}(\mathfrak{A})$ as an ambient space for this.

## 11.2 CC: Convex combinations

**Structure 2** Each $\boldsymbol{c}_\lambda$ above is a *Convex combination (CC)* of $\boldsymbol{c}_0$ and $\boldsymbol{c}_1$ .

**Structure $n$** The $n$-vector generalization of the Convexity condition is as follows.
$$\sum_{i\,=\,1}^{n} \lambda_i\, \boldsymbol{x}_i \ \text{ subject to } \ \sum_{i\,=\,1}^{n} \lambda_i \ = \ 1 \ \text{ and } \ \lambda_i \ \geq \ 0 \ \text{ for each } \ i \ = \ 1 \ \text{ to } \ n\ . \tag{11.2}$$

**Remark 1** This Convex combination condition amalgamates Affineness' 1 equality constraint with conicness' $n$ inequality constraints.

**Remark 2** Compare with the following conditions on the probabilities for $n$ exhaustive and mutually-exclusive events.
$$\sum_{i\,=\,1}^{n} p_i \ = \ 1 \ \text{ and } \ p_i \ \geq \ 0 \ \text{ for each } \ i \ = \ 1 \ \text{ to } \ n$$

**Remark 3** Compare also with partial or fractional notions (Appendix A of Volume 0) in Physics and Chemistry, such as partial moments of inertia or partial pressures.

♠ ◇ ♣

**Exercise 6** Show that subset Convexity can be reformulated as follows. A subset of a vector space is Convex iff it contains every CC of its vectors.

**Remark 4** This means that ascribing primality to combinations (our Introduction's choice in considering linear functions) can however miss out on some alternatives.

**Exercise 7** Referring to Section *not-yet-up's notions if needs be, what relation do CCs bear to mutually-exclusive collectively-exhaustive probabilities?

## 11.3 CS: Convex span

**Definition 1** Suppose that we are given a subset of points
$$\mathfrak{S} \ = \ \{\boldsymbol{s}_i\}_{i\,\in\,\mathfrak{I}} \ \subseteq \ \mathfrak{A}\ .$$
Then the corresponding *Convex-combinational Convex span* is given by the following.
$$\text{CCCS}(\mathfrak{S}) \ := \ \left\{ \sum_{i\,=\,1} \lambda_i\, \boldsymbol{s}_i \ \middle| \ \boldsymbol{s}_i \ \in \ \mathfrak{S}\ , \ \ \lambda_i \ \in \ [\,0,\,1\,] \ : \ \ \sum_{i\,=\,1} \lambda_i \ = \ 1 \right\}\ .$$

**Definition 1′** The *minimal Convex span* $\text{MCS}(\mathfrak{S})$ of $\mathfrak{S}$ is the smallest Convex set containing $\mathfrak{S}$ . I.e.
$$\text{MCS}(\mathfrak{S}) \ := \ \min_{\mathfrak{f}\ \subseteq\ \mathfrak{e}\ \leq\ \mathfrak{a}} \mathfrak{C}\ .$$

**Definition 1″** The *intersectional Convex span* of $\mathfrak{S}$ is the intersection of all the subsets of $\mathfrak{A}$ that contain $\mathfrak{S}$ . I.e.

$$\mathrm{ICS}(\mathfrak{S}) \;=\; \bigcap_{\mathfrak{f} \subseteq \mathfrak{C} \leq \mathfrak{a}}^{\mathfrak{C}} .$$

**Remark 1** We are minimizing, and then instead intersecting, over $\mathfrak{C}$ for a fixed subset $\mathfrak{S}$ in a fixed $\mathfrak{a}$. This can furthermore be identified as minimizing, and then intersecting, *over* the totality of $\mathfrak{C} \leq \mathfrak{a}$. I.e. over $\mathfrak{CS}\mathrm{ubset}(\mathfrak{a})$.

**Exercise 8$^+$** Show that the last two definitions each uniquely define a Convex subset. And then that the above three definitions coincide.

**Naming Remark 1** To celebrate, let us henceforth refer to any of the above three notions as just the *Convex span (CS)*, commonly alias *Convex hull*, which we denote by $\mathfrak{AS}\mathrm{ed}(\mathfrak{S})$.

♠ ◇ ♣

**Example 0** If $\mathfrak{S}$ is Convex, its CS is $\mathfrak{S}$ itself.

**Example 1** The CS of $n$ distinct points in general position is the Convex polytope formed by those points.

**Project 1** What happens if we drop the words 'in general position'?

**Pointer C.0** *Convex polytopes* are particularly interesting, and much theory has been worked out for them (see [13, 14, 18]. Let us also warn the Reader that some authors take *polytope* to mean 'bounded polytope', while others take it to mean 'Convex polytope'. As such, you should always carefully check each source's definition.

## 11.4 CB: Convex bases

**Structure 1** We are now free once again to use the

$$(\,\text{basis}\,) \;=\; (\,\text{minimal spanning set}\,)$$

conceptualization so as to have a notion of Convex basis as well.

**Exercise 9** Show that the centre of mass of a set of points in $\mathbb{R}^n$ must lie within the CS of those points.

**Exercise 10$^+$** Let $\mathfrak{P}\mathrm{erm}(n)$ be the arena of all the $n \times n$ permutation matrices. Investigate which matrices belong to

$$\mathfrak{AS}\mathrm{ed}(\,\mathfrak{P}\mathrm{erm}(n)\,) .$$

Conjecture where you have seen before the matrices forming this new space.

**Exercise 11** Show that each of the current Volume's notions of span must produce connected sets.

[You may assume that connected $\Leftrightarrow$ path-connected in subsets of $\mathbb{R}^n$.]

**Exercise 12** Given a set of points $\mathfrak{Y} \subseteq \mathbb{R}^d$, suppose that every point resides within $\mathrm{CSed}(\mathfrak{Y})$. Reappraise this situation as a notion of (in)dependence. Are there corresponding reformulations of LI, AI and CiI?

## 11.5 Convex dimension $\mathbf{dim_C}$

**Worked Example 4** Is there a meaningful and primary notion of Convex dimension?

**Answer 4** There is a meaningful such, but this

$$\dim_{\mathrm{C}} \ = \ \dim \left( \mathfrak{AS}\mathrm{ed} \right) \ .$$

So by

$$\mathrm{CS} \ = \ \mathrm{AS} \ ,$$

$$\dim_{\mathrm{C}} \ = \ \dim_{\mathrm{A}} \ = \ \dim_{\mathrm{L}} \ .$$

**Project 2** Investigate which of the Restriction, Extension and Replacement Lemmas of Linear Algebra have Convex counterparts.

**Project 3** Investigate the smallest few CC, CI, CSing, CSed and CB arenas.

## 11.6 Pointers on Convex sets themselves

**Pointer C.1** Since quite a lot of concepts and computations are simpler for Convex sets, one quite often elects to work with these when available.

**Project 4** Suppose that we are given two subsets $\mathfrak{Y}$ and $\mathfrak{Z}$ of $\mathbb{R}^n$ . Then their *Minkowski sum* [21]

$$\mathfrak{Y} \ +_{\mathrm{M}} \ \mathfrak{Z} \ := \ \{ y \ + \ z \, | \, y \ \in \ \mathfrak{Y} \ , \ \ z \ \in \ \mathfrak{Z} \} \ . \tag{11.3}$$

a) What is the Minkowski sum of two points? Of two line segments? Of a line segment and a triangle? Of two triangles?

b) Can you find two Convex subsets of $\mathbb{R}^n$ whose Minkowski sum is not Convex?

[Minkowski sums play a significant part in the theory of Convex subsets.]

**Exercise 13$^+$** a) Prove the following two statements.

i) Let $\mathfrak{C}$ be a $c$-dimensional subset of $\mathbb{R}^n$ , and let $\boldsymbol{c} \ \in \ \mathfrak{AS}\mathrm{ed}(\mathfrak{C})$ . Then $\boldsymbol{c}$ can be expressed as a CC of $\leq \ \#c \ = \ c \ + \ 1$ points of $\mathfrak{C}$ . We consequently need less points to represent general points than one might elsewise guess.

ii) Suppose that $\boldsymbol{a}_i \ \in \ \mathbb{R}^n$ for $i \ = \ 1$ to $r \ \geq \ \#r \ \geq \ n \ + \ 2$ . Then the set $\{ 1, \ldots , r \}$ can be bi-partitioned into $\mathfrak{J}_1$ and $\mathfrak{J}_2$ such that the following holds.

$$\bigcap_{a \ = \ 1}^{2} \mathfrak{AS}\mathrm{ed} \left( \{ a_j \}_{j \ \in \ \mathfrak{J}_a} \right) \ \neq \ \emptyset \ .$$

b) Deduce that quadrilaterals are not necessarily Convex.

c) Show that the Theorems' bounds on $\#c$ and $\#r$ are tight.

**Pointer C.2** The above gives a start on Convexity turning out to be useful in Combinatorial Geometry. Volume 3 further fuels this by bringing in *Helly's Theorem*; see also [19].

## 11.7 Convex-conic (CCi) spaces

**Exercise 14** A *Convex-conic combination* (*CCiC*) is a linear combination subject to both of the following.

i) The Convex restriction

$$\sum_{i=1}^{n} \lambda_i \;=\; 1 \;\text{ with each }\; \lambda_i \;\in\; [0,\,1]\;.$$

ii) The conic restriction

$$\lambda_i \;\geq\; 0$$

What happens here notions of (in)dependence, span, basis and dimension (CCiI, CCiD, CCiS, CCiB, $\dim_{\text{CCi}}$ ?

**Motivation 1** Convex-conic spaces turn out to be useful in Functional Analysis [15].

**Exercise 15** What happens to Exercise **??**.11's answer if we additionally require such subsets to be Convex?

**Exercise 16** Account for how subsets form the poset in Fig 1.3.a) under subset inclusion. And yet the corresponding notions of combination under the amount of constraints being applied form the distinct one in Subfigure b).

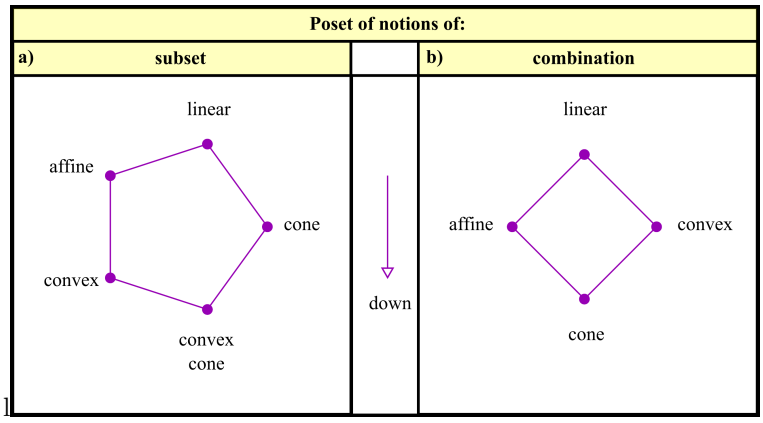


Figure 11.3:

**Project 5**$^{+}$ Investigate which of the Restriction, Extension and Replacement Lemmas of Linear Algebra have CCi counterparts.

**Project 6**$^{+}$ Investigate the smallest few CiCC, CiCI, CiCSing, CiCSed and CiCB arenas.

[Treat the above two Questions as bonus questions, which you might try if you are particularly interested in Functional Analysis or Optimization!]

# 1.B. Convex functions

## 11.8 Defining Convex and concave functions

**Definition 1** A function $f : \mathfrak{V} \longrightarrow \mathbb{R}$ is *Convex* if the following holds.

$$f(\lambda \boldsymbol{a} + (1-\lambda)\boldsymbol{b}) \leq \lambda f(\boldsymbol{a}) + (1-\lambda) f(\boldsymbol{b}) \ \forall \lambda \in (0,1) \text{ and } \forall \boldsymbol{a} \neq \boldsymbol{b} \in \mathfrak{V} \quad (11.4)$$

**Definition 2** $f$ is *concave* if $-f$ is Convex.

**Exercise 17** Check that this reflection-trick definition reverses the sign of (1.4)'s inequality.

**Exercise 18** a) Functions certainly need not be either Convex or concave. Show that all of a generic cubic such as $x(x-1)(x-2)$, $\sin x$ and $\text{cosec} x$ are neither.

b) There is however a sense in which Convexity and concavity are local properties of functions. Find closed intervals on which each of a)'s examples are Convex and others on which they are concave.

**Exercise 19** For functions $\mathbb{R} \longrightarrow \mathbb{R}$, show that vertically or horizontally translations make no difference to their Convexity properties. Nor does stretching the function parallel to the $y$-axis, or squeezing it parallel to the $x$-axis.

## 11.9 Inter-relating Convex regions and Convex functions

**Worked Example 5** i) Show that Convex and concave functions have corresponding Convex-set regions.

ii) Show that Convex subsets correspond to a subcase of Convex functions.

**Answer 1** i) Introduce the notion of *epigraph* of a function. This is a 'graph' in a geometrical rather than combinatorial sense. I.e. the

$$(x, y) = (x, f(x))$$

plot. The epigraph itself is the region above $f(x)$ in the $(x, y)$ plane. I.e.

$$\mathfrak{Epi}(f) := \{(x, k) \in \mathfrak{X} \times \mathbb{R} \,|\, f(x) \leq k\} .$$

Claim: $f$ is Convex as a function iff $\mathfrak{Epi}(f)$ is as a subset. Probe with two arbitrary points

$$(x_0, k_0), \quad (x_1, k_1), \quad \text{both} \in \mathfrak{Epi}(f) .$$

Then for $\mathfrak{Epi}(f)$ a Convex subset, for each $\lambda \in (0,1)$, the following holds.

$$(x_\lambda, k_\lambda) := \lambda (x_0, k_0) + (1-\lambda)(x_1, k_1) \in \mathfrak{Epi}(f) \text{ too} .$$

$$\Leftrightarrow f(x_i) \leq k_i, \quad i = 0, 1 \Rightarrow f(x_\lambda) \leq k_\lambda \text{ for each } \lambda \in (0,1)$$

I.e. $f$ meets the definition of Convex function.

ii) Convex subsets $\mathfrak{C}$ can be characterized by indicator functions $\iota_{\mathfrak{C}}$ (Chapter 6 of Volume 0). Since these are indeed functions,

$$(\text{Convex subsets}) \subseteq (\text{functions}) .$$

This is furthermore clearly a strict inclusion $\subset$, since any function somewhere taking the value 2 is not an indicator function.

**Exercise 20** The above working is not complete since i) has not covered the concave case. Modify i)'s working to deal with concavity via introducing the following notion of *hypograph*.

$$\mathfrak{Hypo}(f) := \{(x, k) \in \mathfrak{X} \times \mathbb{R} \,|\, f(x) \geq k\} .$$

## 11.10 A brief etymology and history

**Naming Remark 2** 'Convex' derives from 'convexus', meaning 'arched' or 'vaulted' in Latin. This is furthermore part of the verb 'convehere'. Which is a composite of, firstly, the same 'vehere' meaning 'to transport' that we already encountered in discussing the word 'vector'. Secondly, of 'cum' meaning 'together'. So 'convexus' means 'brought together', along the lines of an arch holding up a $2$-$d$ piece of wall, or a vault holding up a 2-surface roof within $3$-$d$ space. Thereby understanding this name indeed points us to various examples of Convex subsets (including the region under the Roman arch of Fig 1.2.10).

'Concave' means 'hollow together', now deriving from 'cavus': Latin for 'hollow'.

**Historical Remark** Convexity dates at least as far back as Archimedes [1, 3]. Systematic treatment of Convexity only started in the mid 20th Century [9, 10, 12, 11]; Cauchy [3], Brunn [4] and Minkowski [5, 6, 7] were 'early birds' prior to this. Euler missed out by a small margin a century earlier [2]. For reviews of Convexity, see e.g. [21, 16, 19].

**Naming Remark 3** 'Epigraph' means 'graph above', by use of the Greek prefix '$\epsilon\pi\iota$' (compare 'epidermis', and 'epicentre'). While 'hypograph' means 'graph below', by use of the opposite Greek prefix '$\upsilon\pi o$ (compare 'hypodermic' and 'hypotenuse': the side 'stretched under' the right angle).

## 11.11 Line-segment gradients

**Remark 1** The *gradient* of our line segments is the dilatational quantity (i.e. ratio of differences)

$$g_f(b,\, a) \;=\; \frac{f(b) \;-\; f(a)}{b \;-\; a} \;.$$

**Remark 2** Our line segment is then given by the equation

$$g_f(b,\, a) \;=\; g_f(x,\, a)$$

since the gradient is the same for all points on the line segment. For

$$x \;\in\; [\, c,\, d\,] \;,$$

so as to restrict from the whole line to the line segment itself.

Explicitly,

$$\frac{f(b) \;-\; f(a)}{b \;-\; a} \;=\; \frac{f(x) \;-\; f(a)}{x \;-\; a} \;.$$

This rearranges to the standard form for the equation of a line in $2$-$d$ as follows.

$$y \;=\; f(x) \;=\; \frac{f(b) \;-\; f(a)}{b \;-\; a}\, x \;+\; f(a) \;.$$

**Remark 3** This lies above-or-on our original function $f$ if

$$g_f(b,\, a) \;\geq\; g_f(x,\, a) \;.$$

I.e.

$$\frac{f(b) \;-\; f(a)}{b \;-\; a} \;\geq\; \frac{f(x) \;-\; f(a)}{x \;-\; a} \;.$$

This provides an equivalent definition of Convexity [20].

**Remark 4** Keeping $a$ fixed, we can bring in $b$ closer and closer (Fig 1.4). Then parametrizing by

$$b \;=\; a \;+\; h \;,$$

for a locally-convex piece of a function, we obtain the following.

$$\frac{f(a + h_1) - f(a)}{b - a} < \frac{f(a + h_2) - f(a)}{x - a} \quad \text{for} \quad h_1 < h_2 .$$

This is a position to which we shall return once equipped with Calculus.

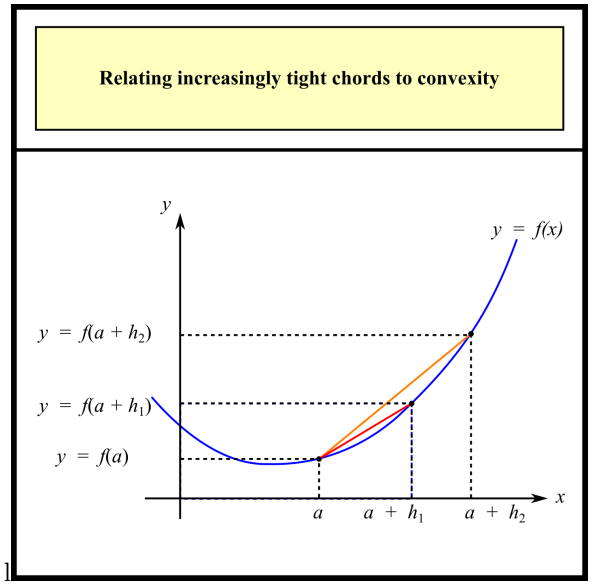


Figure 11.4:

**Exercise 21** A function $f : \mathfrak{V} \longrightarrow \mathbb{R}$ is *strictly convex* if the following holds.

$$f(\lambda \boldsymbol{a} + (1 - \lambda)\boldsymbol{b}) < \lambda f(\boldsymbol{a}) + (1 - \lambda) f(\boldsymbol{b}) \ \forall \lambda \in (0, 1) \text{ and } \forall \boldsymbol{a} \neq \boldsymbol{b} \in \mathfrak{V} \quad (11.5)$$

a) Set up strict convexity for a region, rather.

b) Show that our 2-way correspondence between region and function Convexity carries over to strict Convexity.

c) Give examples of strictly convex and yet not Convex subsets of $\mathbb{R}^2$ .

## 11.12 A start on motivation for Convexity

**Pointer C.3** Calculus provides further useful diagnostics for Convexity; see Sec **??**.

**Pointer C.4** Some applications of Convexity proceed moreover via its provision of useful inequalities, of which these next two Sections provide a first taste. Ties to Probability Theory are evident from the first of these sections.

**Pointer C.5** Exercise 22 and **??**.31 's functions furthermore have special significance as entropies in Thermodynamics, which, upon flipping sign, are also notions of information. These subjects then further benefit from Convexity properties starting with those derived in these Exercises.

**Pointer C.6** Convexity is a simplifying feature for Optimization problems to have, at both the Pre-Calculus (Chapter **??**) and Calculus levels (Sec *-not-up yet).

## 11.13 Jensen's inequality for Convex functions

**Definition 1** For a discrete random variable, the expectation takes the form

$$\mathbb{E}(x) \;=\; \sum_{i\,=\,1}^{n} p_i\, f\,(x_i) \;\;.$$

For exhaustive and mutually-exclusive probabilities $p_i$ and weighting functions $f(x_i)$ .

**Theorem 1 (Jensen's inequality)** [8] Let $f \;:\; (\,a,\,b\,) \longrightarrow \mathbb{R}$ be a Convex function. Then
i)

$$\mathbb{E}(f(x)) \;\geq\; f(\mathbb{E}(f(x))) \;\;.$$

ii) Equality occurs iff

$$x_1 \;=\; x_2 \;=\; \ldots \;=\; x_n \;\;.$$

Proof (Discrete case only.) i) Strongly induct on $n$ .

Base step. For $n \;=\; 1$ , there is nothing to prove. For $n \;=\; 2$ , the result follows from the definition of Convex function.

Induction step. Assume the result holds up to $n \;-\; 1$ . Let $x_i \;\in\; (\,a,\,b\,)$ for $i \;=\; 1$ to $n$ , and $p_i \;\in\; [\,0,\,1\,]$ such that

$$\sum_{i\,=\,1}^{n} p_i \;=\; 1 \;\;.$$

Single out one $p_i$ to treat differently, say $p_1$ . We can assume that $p_1 \;\neq\; 1$. For if it was, then our result would hold in the trivial equality form $f\,(x_1) \;=\; f\,(x_1)$ . For $j \;=\; 2$ to $n$ , we can thus licitly set

$$p'_j \;=\; \frac{p_j}{1 \;-\; p_1} \;\;. \tag{11.6}$$

Also

$$\sum_{j\,=\,2}^{n} p_j \;=\; \sum_{i\,=\,1}^{n} p_i \;-\; p_1 \;=\; 1 \;-\; p_1 \;\;. \tag{11.7}$$

And so

$$\sum_{j\,=\,2}^{n} p'_j \;\;=\;\; \sum_{j\,=\,2}^{n} \frac{p_j}{1 \;-\; p_1} \;\;=\;\; \frac{1}{1 \;-\; p_1} \sum_{j\,=\,2}^{n} p_j \;\;=\;\; \frac{1 \;-\; p_1}{1 \;-\; p_1} \;\;=\;\; 1 \;\;.$$

This is by (1.6), then pulling out a common factor, and then (1.7).

Finally,

$$\mathbb{E}(f(x)) \;=\; \sum_{i\,=\,1}^{n} p_i\, f\,(x_i) \;=\; p_1\, f\,(x_1) \;+\; \sum_{j\,=\,2}^{n} p_j\, f\,(x_j)$$

$$=\; p_1\, f\,(x_1) \;+\; (\,1 \;-\; p_1\,) \sum_{j\,=\,2}^{n} p'_j\, f\,(x_j) \;\;\geq\;\; p_1\, f\,(x_1) \;+\; (\,1 \;-\; p_1\,)\, f\left( \sum_{j\,=\,2}^{n} p'_j\, x_j \right)$$

$$\geq\;\; f\left( p_1\, f\,(x_1) \;+\; (\,1 \;-\; p_1\,)\, f\left( \sum_{j\,=\,2}^{n} p_j\, x_j \right) \right) \;\;=\;\; f\left( \sum_{i\,=\,1}^{n} p_i\, x_i \right) \;\;=\;\; f(\mathbb{E}(x)) \;\;.$$

The first step is by the definition of expectation. The second step splits the sum. The third makes two uses of the induction hypothesis: for $2$ and for $n \;-\; 1$ . The fourth is by the definition of Convex function. The fifth re-forms the sum. The final step is by the definition of expectation in reverse.

ii) Suppose that $f$ is strictly convex. If some $x_i \neq x_j$, then the following step enters our working.

$$\sum_{j=2} {p_j}' f(x_j) \quad > \quad f\left( \sum_{i=1}^{n} p_i f(x_i) \right) .$$

Jensen's inequality must thus be strict. Finally apply the contrapositive. □

## 11.14 Exercises on inequalities from Convexity

**Exercise 22** i) Show that $f(x) = -\ln x$ is Convex.

ii) Specializing Jensen's inequality to this case, deduce the *AM-GM* (arithmetic mean–geometric mean) *inequality*

$$\left( \prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} \quad \leq \quad \frac{1}{n} \sum_{i=1}^{n} x_i .$$

**Exercise 23** A function $f : \mathfrak{V} \longrightarrow \mathbb{R}$ is *log-convex* if $\log f$ is Convex.

i) Show that for such $f$,

$$f(\lambda \boldsymbol{a} + (1 - \lambda)\boldsymbol{b}) \leq f(\boldsymbol{a})^{\lambda} f(\boldsymbol{b})^{1-\lambda} .$$

ii) Which extra caveat arises in defining *log-concavity*? Show also that $f$ is log-convex iff $1/f$ is log-concave.

iii) For which $\alpha \in \mathbb{R}$ are $x^{\alpha}$ and $\exp(\alpha x)$ log-convex?

iv) Show that for positive-definite matrices $\boldsymbol{M}$, the following are log-concave. $\det \boldsymbol{M}$ and the ratio of invaraints

$$R(\boldsymbol{M}) := \frac{\det \boldsymbol{M}}{\operatorname{tr} \boldsymbol{M}} .$$

v) Show that log-convex $\Rightarrow$ Convex, and also develop a concave counterpart of this result.

**Exercise 24** Show that if $f$ is log-concave, then

$$f\left( \frac{x + y}{2} \right) \quad \geq \quad \sqrt{f(x) f(y)} .$$

**Exercise 25** i) From $\ln x$ being concave, deduce *Young's inequality*

$$x y \quad \leq \quad \frac{x^p}{p} + \frac{x^q}{q}$$

for

$$\frac{1}{p} + \frac{1}{q} = 1 .$$

ii)$^{+}$ Deduce *Hölder's inequality*

$$\sum_{i=1}^{n} |x_i y_i| \quad \leq \quad \left( \sum_{i=1}^{n} {x_i}^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^{n} {y_i}^q \right)^{\frac{1}{q}} .$$

iii)$^{+}$ Deduce *Minkowski's inequality*

$$\left( \sum_{i=1}^{n} |x_i + y_i|^p \right)^{\frac{1}{p}} \quad \leq \quad \left( \sum_{i=1}^{n} {x_i}^p \right)^{\frac{1}{p}} + \left( \sum_{i=1}^{n} {y_i}^p \right)^{\frac{1}{p}} .$$

**Remark 1** See Sec *not-up-yet for continuum functions and probabilities' counterparts of the above proof and examples.

## 11.15 Aside: commutators as measures of non-commutativity

**Definition 1** The (infinitesimal algebraic) *commutator* of $A$ and $B$ is

$$[\, A, B \,] \;:=\; A\, B \;-\; B\, A \;.$$

**Remark 1** While this structure has many more advanced and well-known applications in Geometry, Lie Theory and Physics (Chapter **??** and Volumes 3 and 6-9), for now let us present some Linear Algebra and Convexity applications.

**Application 1** We can now rephrase (**??**) as

$$\left[ \sum_{o \,=\, 1}^{O} \;,\; \dim \right] \mathfrak{V}_o \;=\; 0 \;.$$

**Application 2 and Naming Remark 4** The *Jensenator*

$$J \;:=\; [\, \mathbb{E}, f \,] \;.$$

Jensen's inequality can then be rephrased as

$$J \;\geq\; 0 \tag{11.8}$$

On the one hand, (1.8) is a standard repackaging. On the other hand, it has hitherto been called the *Jensen gap*. And yet *Jensenator* identifies it as one of many geometrical measures of deviation that are commutators.

For a general function, $J$ can be of any sign. But for a Convex function, it must be non-negative. And this is useful in developing, and in working with, Probability Theory.

**Remark 2** In contrast, Application 1 is a simpler compatibility criterion: no deviation upon re-ordering inputs.

**Remark 3** Definition 1 requires $A$ and $B$ to belong to spaces supporting notions of subtraction and of some kind of product or composition of operators. The above definition is not for now intended to provide a sharp formulation. After all, we are currently using commutators *conceptually* as measures of non-commutation (for products) or of exchanging operations (for compositions). The above two applications are compositional.

# Bibliography


[1] Archimedes of Syracuse, “On the Sphere and the Cylinder";

“On the Equilibrium of Plane Figures"

For an English translation, see e.g. J.L. Heath, *The Works of Archimedes* (Dover, N.Y. 1912).

[2] L. Euler implicitly assumed Convexity in treating $V - E + F = 2$ in 1752. He then failed to take into account that his re-triangulation induction step does not necessarily preserve Convexity. By which this work contains a rare example of an error made by Euler, and an error for a subsequently-interesting reason at that.

[3] According to W. Fenchel [**?**], A. Cauchy wrote on Convexity in 1841 and 1850.

[4] H. Brunn, “Über Ovale und Eiflache" (On Ovals and Egg Shapes) Ph.D Thesis (Munich, 1887);

“Über Curven und Wendepunkte" (On Curves and Turning Points) Habilitation Thesis (Munich, 1889).

[5] H. Minkowski, “Uber die positiven quadratischen Formen und über kettenbruchähnliche Algoritmen”, p 243 of *Gesammelte Abhandlungen* **1** (Teubner, Leipzig 1911); the original paper dates to 1891.

[6] H. Minkowski, *Geometrie der Zahlen* (Teubner, Leipzig 1896). For an English translation, see e.g. *The Geometry of Numbers* (Chelsea, New York 1953).

[7] H. Minkowski, “On a generalization of Surface Area" (1901).

[8] J.L.W.V. Jensen, *Sur les Fonctions Convexes et les Inégalités entre les Valeurs Moyennes*, Acta Math. **30** 175 (1906).

[9] T. Bonnesen and W. Fenchel, *Theory of Convex Bodies* (Springer, Berlin 1934) (in German). For an English translation, see e.g. (Moscow, Idaho 1987).

[10] J. von Neumann, Collected works. Vol II. pp. 508–527 dating to 1935.

[11] M. Krein and D. Milman, "On extreme points of regular convex sets", Studia Mathematica **9** 133 (1940).

[12] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (P.U.P, Princeton N.J. 1944).

[13] H.S.M. Coxeter, *Regular Polytopes* (Methuen 1947; Dover, New York 1971).

[14] B. Grünbaum, *Convex Polytopes* (Wiley, N.Y. 1967). This first edition was historically significant. Yet the modern Reader would do better to consult the second edition ed. V. Kaibel, V. Klee, and G.M. Ziegler (Springer, N.Y. 2003).

[15] R.T. Rockafellar *Convex Analysis*, (P.U.P., Princeton N.J. 1970).

[16] S.R. Lay, *Convex Sets and their Applications* (Wiley, New York 1982).

[17] P.D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, N.Y. 1982)

[18] A. Brondsted, *An Introduction to Convex Polytopes* (Springer, New York 1983).

[19] R.J. Webster, *Convexity*, (O.U.P., Oxford 1994): pp. 1-72 and 193-231 are suitable background reading at this stage. **CE**

[20] M. Spivak, *Calculus* (C.U.P., Cambridge 2006).

[21] N. Lauritzen, *Undergraduate Convexity* (World Scientific, Singapore 2013).