coursera

Convexity: clearing the air

Let us for a moment summarise all we have learned about convexity.

Definition and corrections

Let us address the definition for the single-variate case first (the multi-variate is a simple generalization as we discussed): *the function is convex on the segment* $[a;b]$ *if the function's graph on any subsegment* $[c;d]$ *is below the segment connecting points* $(c, f(c))$ *and* $(d, f(d))$.

We came up with a more algebraic way to write it:

- if we have a segment $[c;d]$, then any point there could be written as $\alpha c + (1-\alpha)d$ where $0 \leq \alpha \leq 1$; why is it so? One way to think about it is the following: imaging a seesaw which starts at c and end at d: if we put a weight α on one side and $1-\alpha$ on the other, then where should we put the foundation (the fulcrum) so it is in balance? Exactly in $\alpha c + (1-\alpha)d$ (it is known in physics as the center of mass). Now try to change α and you will see how we can get the fulcrum to any point of the segment
- the similar thing applies to the two dimensional segment (e.g. between $(c, f(c))$ and $(d, f(d))$) -- we still need to sum them with α and $1-\alpha$ weights
- now all we need is to write an inequality for the word "below" in the definition: $f(\alpha c + (1-\alpha)d) \leq \alpha f(c) + (1-\alpha)f(d)$

Attention: there is a correction in comparison with our