We've seen QUICKSORT $\simeq$ BST-SORT } Randomized versions have same analysis

(stable)

We've seen QUICKSORT $\simeq$ BST-SORT $\Big\}$ Randomized versions have same analysis

(stable)

$E[\text{BST build}] = \Theta(n \log n)$ time

$E[\text{insert node}] = \Theta(\log n)$ time

We've seen QUICKSORT $\simeq$ BST-SORT } Randomized versions have same analysis

(stable)

$E[\text{BST build}] = \Theta(n \log n)$ time

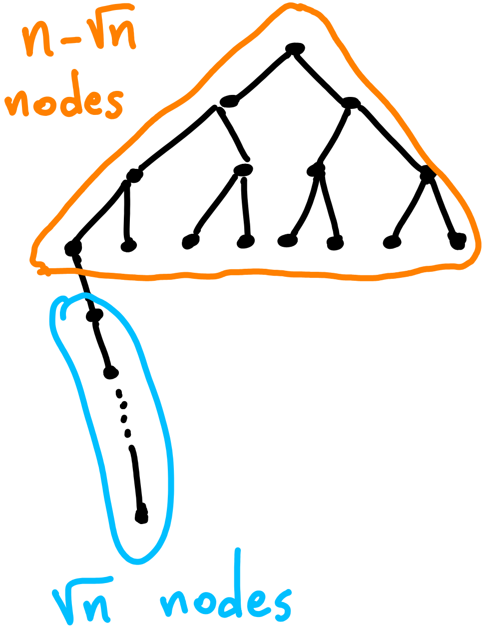
$E[\text{insert node}] = \Theta(\log n)$ time

Now, forget the sorting context of BST.

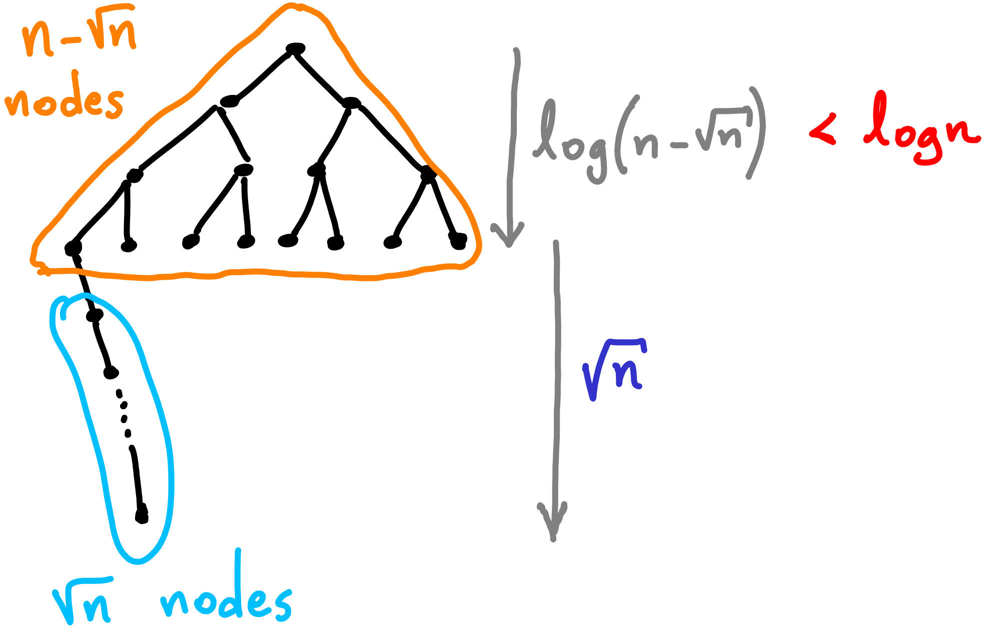
Let's just examine what an average random BST looks like

Intuition : $E[\text{depth}] = \Theta(\log n)$ so it should be $\sim$balanced ?

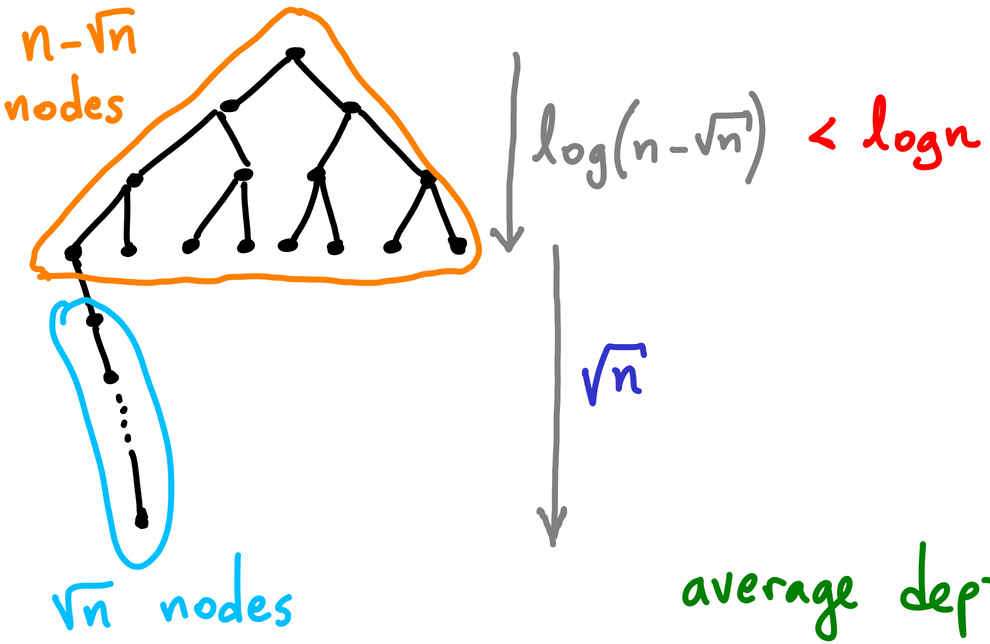
Intuition : $E[\text{depth}] = \Theta(\log n)$ so it should be ~balanced ? NO



Intuition : $E[\text{depth}] = \Theta(\log n)$ so it should be ~balanced ? NO

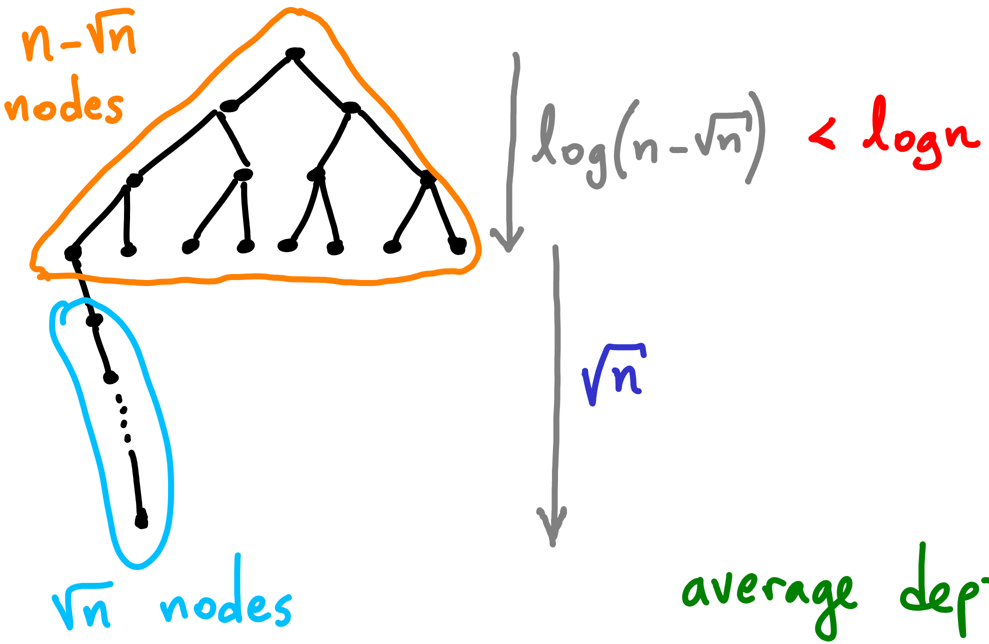


Intuition : $E[\text{depth}] = \Theta(\log n)$ so it should be ~balanced ? NO



$$\text{average depth} < \frac{1}{n}\cdot\Big(\overset{\text{exaggerated \#nodes}}{n}\cdot\overset{\text{exaggerated depth}}{\log n} + \overset{\#\text{nodes}}{\sqrt{n}}\cdot\overbrace{(\sqrt{n}+\log n)}^{\text{exaggerated depth}}\Big)$$

Intuition : $E[\text{depth}] = \Theta(\log n)$ so it should be ~balanced ? NO



$$\text{average depth} < \frac{1}{n}\cdot\Big(\overbrace{n}^{\text{exaggerated \#nodes}}\cdot\overbrace{\log n}^{\text{exaggerated depth}} + \overbrace{\sqrt{n}}^{\text{\#nodes}}\cdot\overbrace{(\sqrt{n}+\log n)}^{\text{exaggerated depth}}\Big)$$

$$= \log n + 1 + \frac{\log n}{\sqrt{n}}$$

$$= O(\log n)$$ so $E[\text{depth}] \not\Rightarrow$ balance

$E[\text{height of random BST}] = O(\log n)$: long PROOF

E[height of random BST] $= O(\log n)$: long PROOF

sketch: (1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$E[\text{height of random BST}] = O(\log n)$: long PROOF

sketch:

(1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

(2) X_n = random variable = height of random tree w/ n nodes

$E[\text{height of random BST}] = O(\log n)$: long PROOF

sketch: (1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

(2) For X_n = random variable = height of random tree w/ n nodes,
prove $E[2^{X_n}] = O(n^3)$

$E[\text{height of random BST}] = O(\log n)$: long PROOF

sketch: (1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

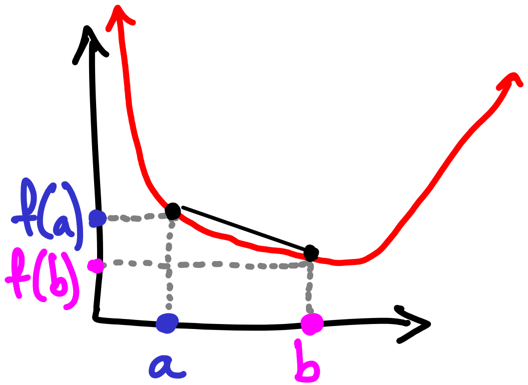
(2) For X_n = random variable = height of random tree w/ n nodes,
prove $E[2^{X_n}] = O(n^3)$

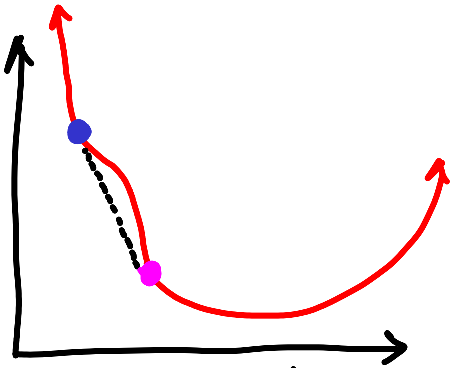
easy- (3) Use (1) & (2): $\underbrace{2^{E[X_n]} \leq E[2^{X_n}]}_{(1)} \Rightarrow E[X_n] = O(\log n^3)$

(1) Jensen's inequality proof & intuition

(1) Jensen's inequality proof & intuition

Convex function f

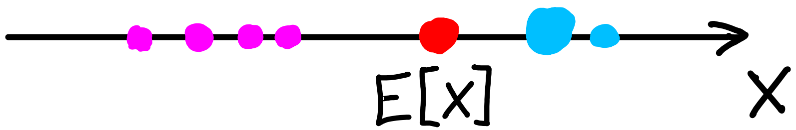




The segment joining any two points $(a, f(a))$ & $(b, f(b))$ must not be below the curve f.

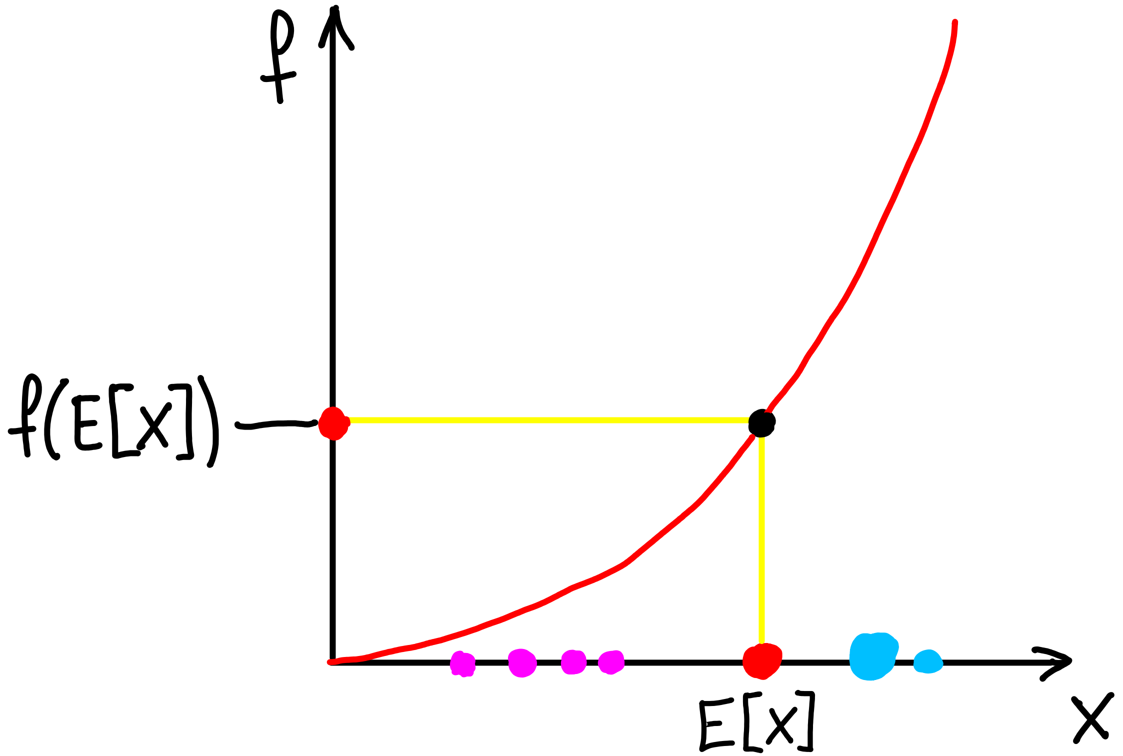
Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$E[X]$ = weighted average of X



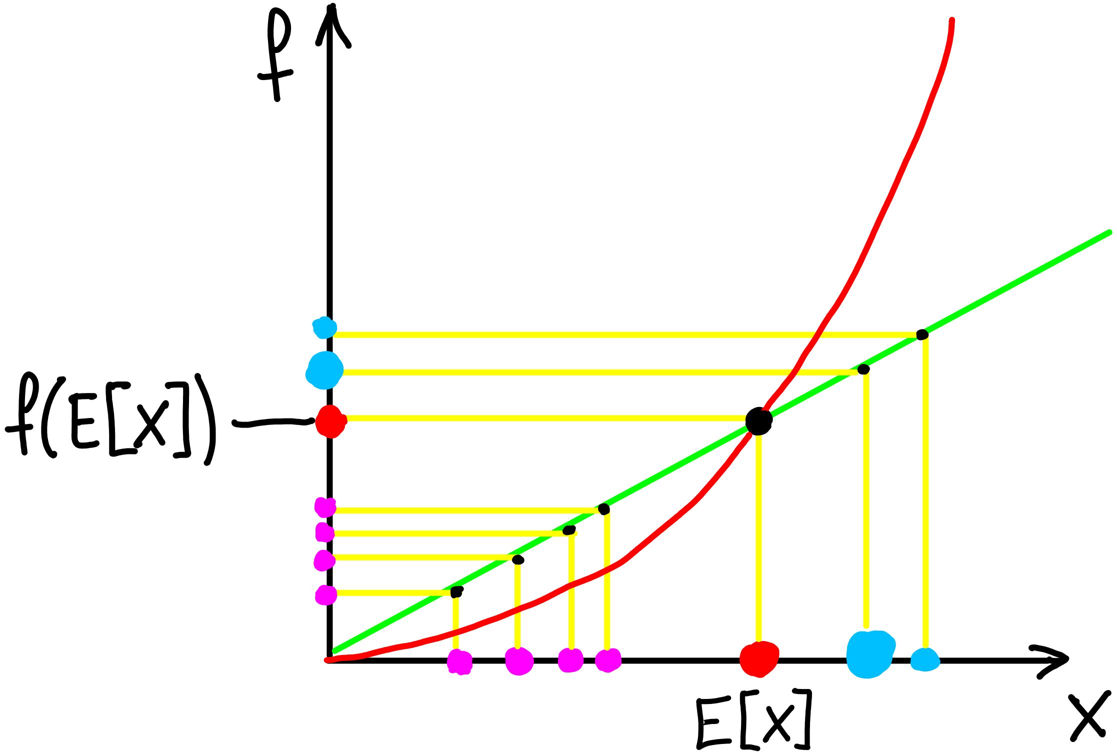
Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$f(E[X]) = f(\text{weighted average of } X)$



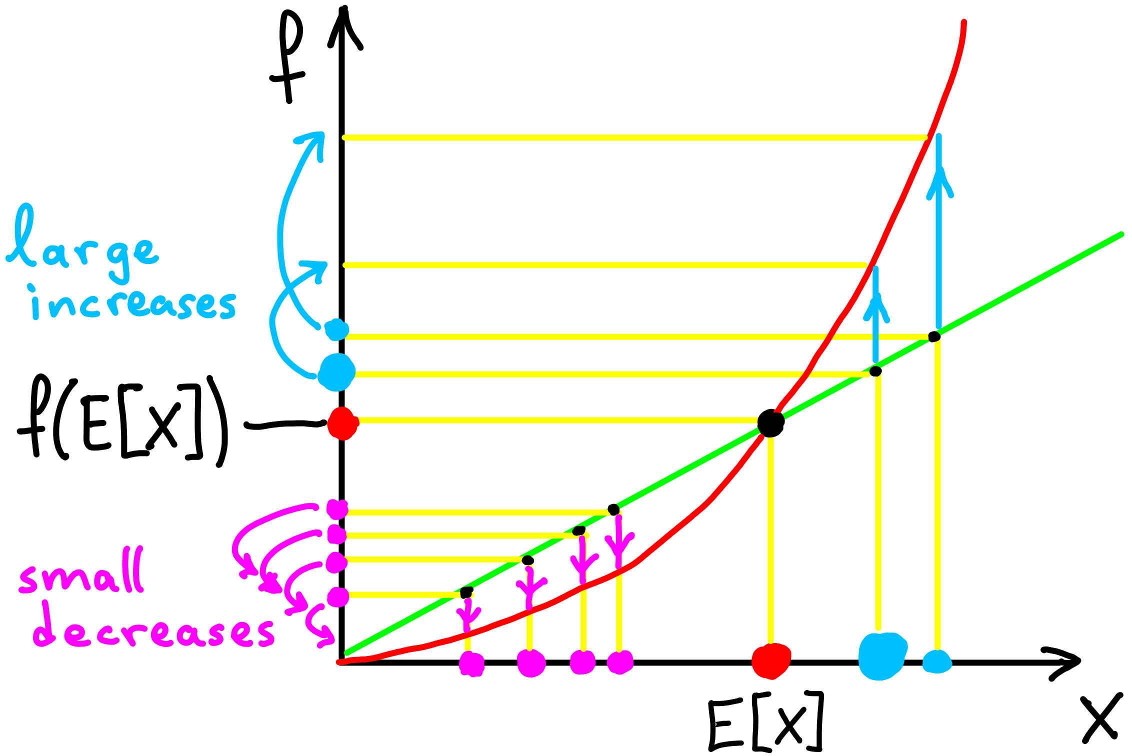
Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$f(E[X]) = f(\text{weighted average of } X)$



Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

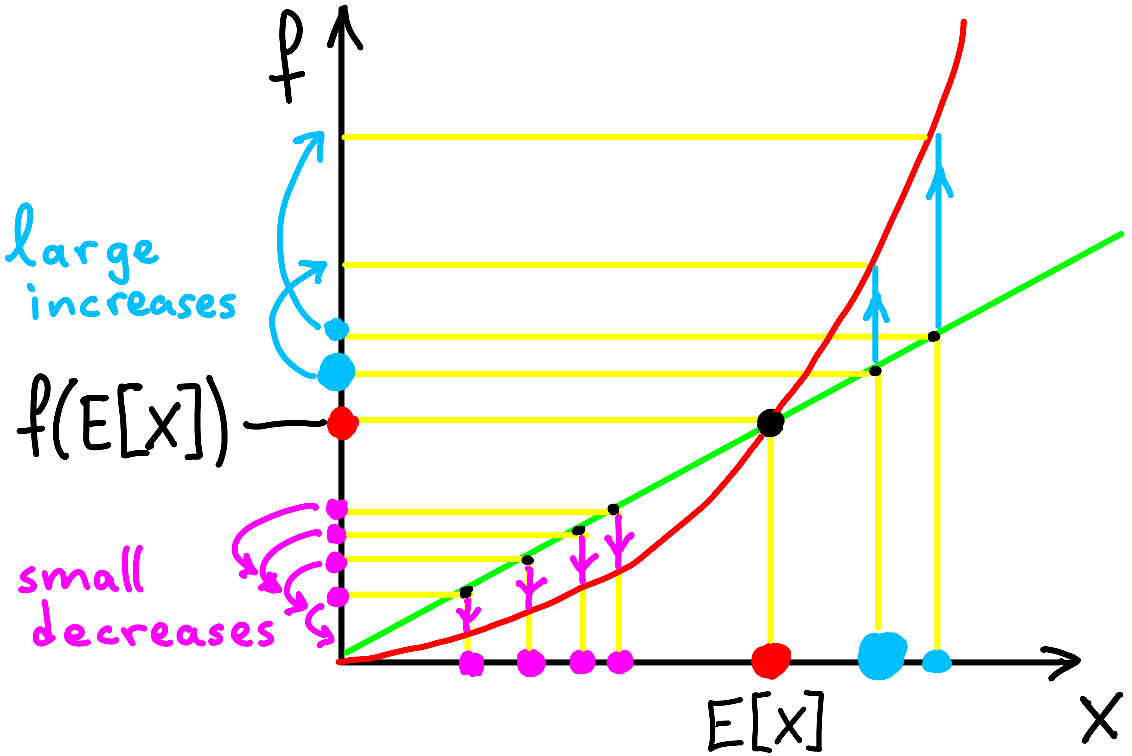
$f(E[X]) = f(\text{weighted average of } X)$



Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$f(E[X]) = f(\text{weighted average of } X)$

$E[f(x)] =$ weighted average height of all $f(X)$

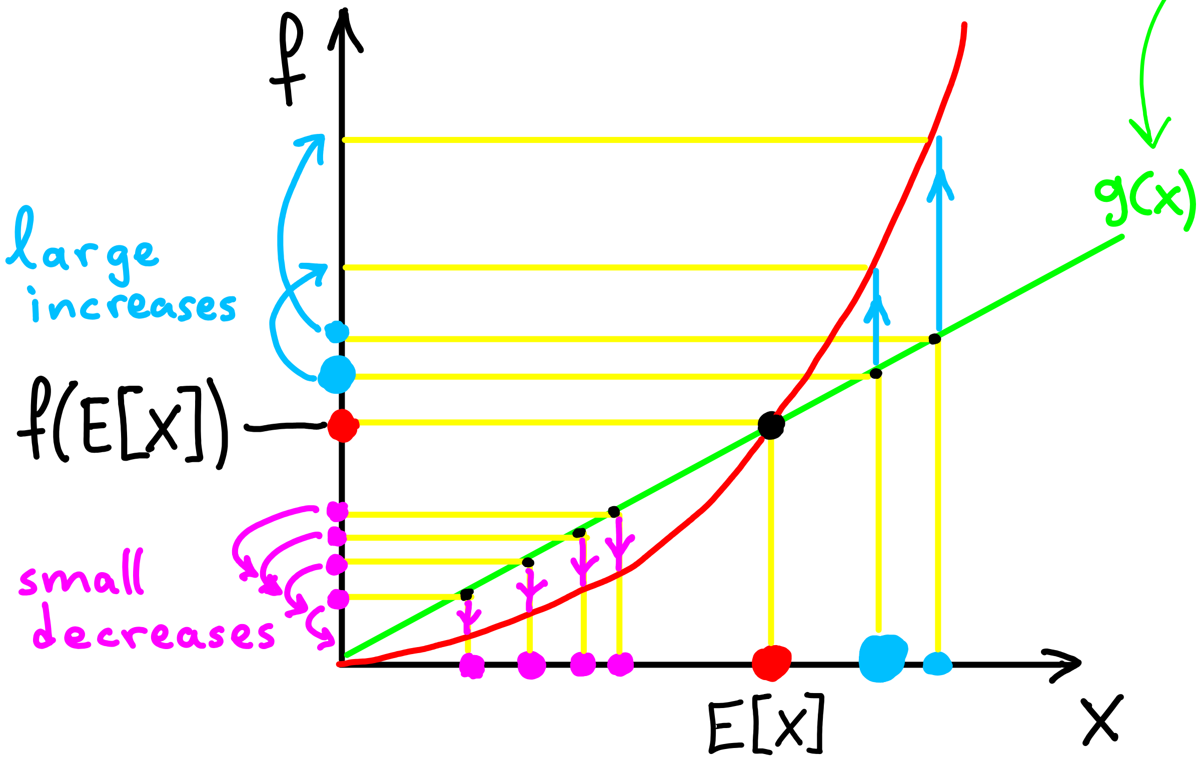


Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$f(E[X]) = f(\text{weighted average of } X)$

$E[f(x)]$ = weighted average height of all $f(X)$

↳ would $= f(E[X])$ if using

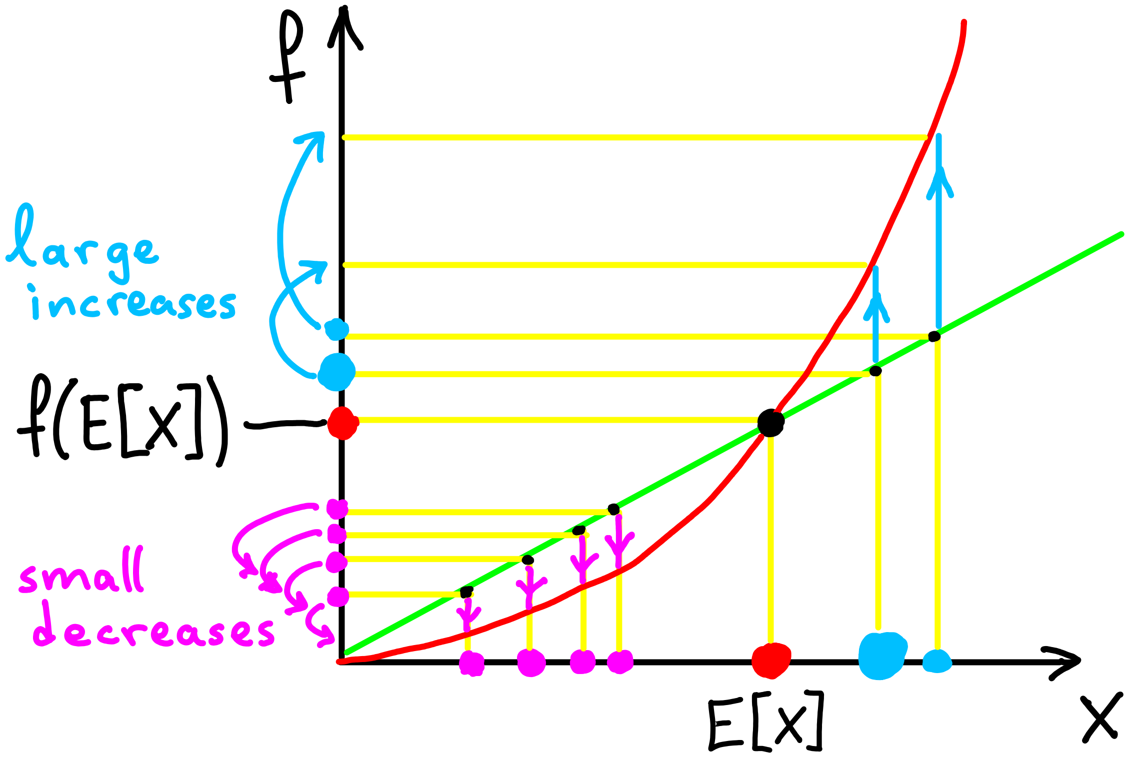


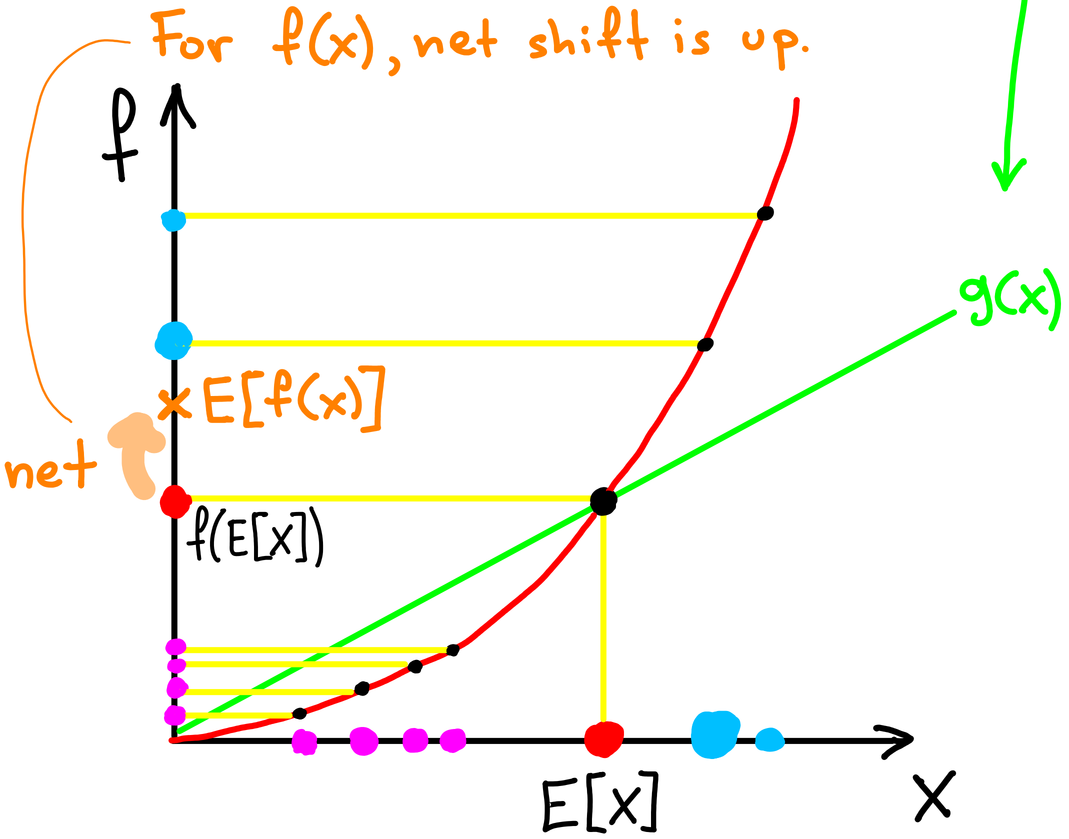
Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$f(E[X]) = f(\text{weighted average of } X)$

$E[f(x)] =$ weighted average height of all $f(X)$

↳ would $= f(E[X])$ if using





More about convexity & linear combinations

Given $x_1\ x_2\ x_3 \ldots x_n \in \mathbb{R}$ (input)

More about convexity & linear combinations

Given $x_1\ x_2\ x_3 \dots x_n \in \mathbb{R}$ (input)

& $\alpha_1\ \alpha_2\ \alpha_3 \dots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$

More about convexity & linear combinations

Given $x_1\ x_2\ x_3 \dots x_n \in \mathbb{R}$ (input)

& $\alpha_1\ \alpha_2\ \alpha_3 \dots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$

then $\min\{x_i\} \leq \sum_{i=1}^{n} \alpha_i x_i \leq \max\{x_i\}$

More about convexity & linear combinations

Given $x_1\ x_2\ x_3 \ldots x_n \in \mathbb{R}$ (input)

& $\alpha_1\ \alpha_2\ \alpha_3 \ldots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$

then $\min\{x_i\} \leq \sum_{i=1}^{n} \alpha_i x_i \leq \max\{x_i\}$

ex: $\alpha_2 = 1$; $\alpha_{\neq 2} = 0 \rightarrow \sum = x_2$

More about convexity & linear combinations

Given $x_1\ x_2\ x_3 \dots x_n \in \mathbb{R}$ (input)

& $\alpha_1\ \alpha_2\ \alpha_3 \dots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$

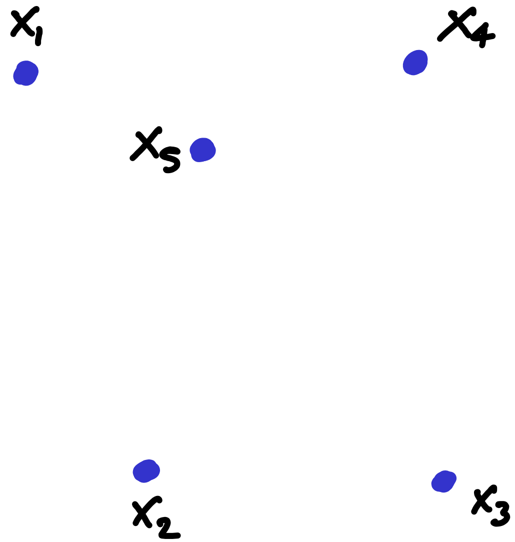
then $\min\{x_i\} \leq \sum_{i=1}^{n} \alpha_i x_i \leq \max\{x_i\}$

ex: $\alpha_2 = 1$; $\alpha_{\neq 2} = 0 \rightarrow \Sigma = x_2$

ex2: $\alpha_i = \frac{1}{n} \rightarrow \Sigma =$ center of mass

More about convexity & linear combinations in 2D

Given $x_1\ x_2\ x_3 \dots x_n \in \mathbb{R}^2$ (input)

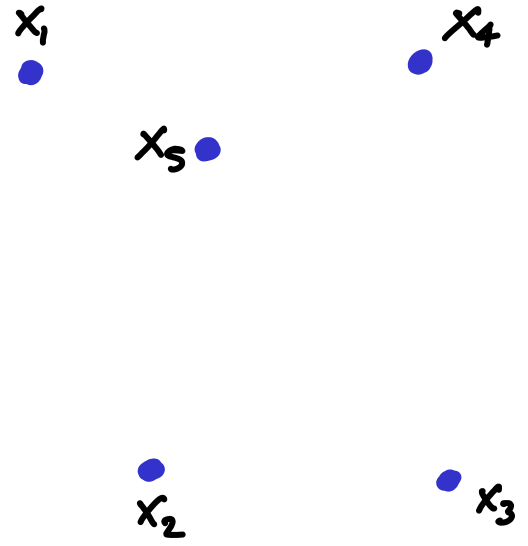


More about convexity & linear combinations in 2D

Given $x_1\ x_2\ x_3 \ldots x_n \in \mathbb{R}^2$ (input)

& $\alpha_1\ \alpha_2\ \alpha_3 \ldots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$



More about convexity & linear combinations in 2D

Given $x_1\ x_2\ x_3 \ldots x_n \in \mathbb{R}^2$ (input)

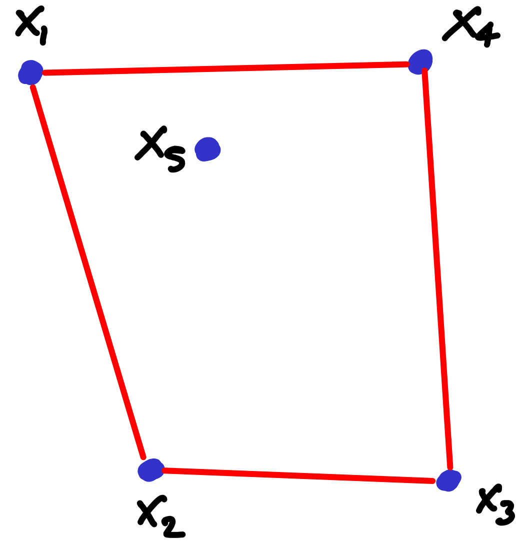
& $\alpha_1\ \alpha_2\ \alpha_3 \ldots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$

then $\cancel{\min\{x_i\}} \leq \sum_{i=1}^{n} \alpha_i x_i \leq \cancel{\max\{x_i\}}$

↳ is inside CONVEX HULL :

(a definition of "extreme")



More about convexity & linear combinations in 2D

Given $x_1\ x_2\ x_3 \dots x_n \in \mathbb{R}^2$ (input)

& $\alpha_1\ \alpha_2\ \alpha_3 \dots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

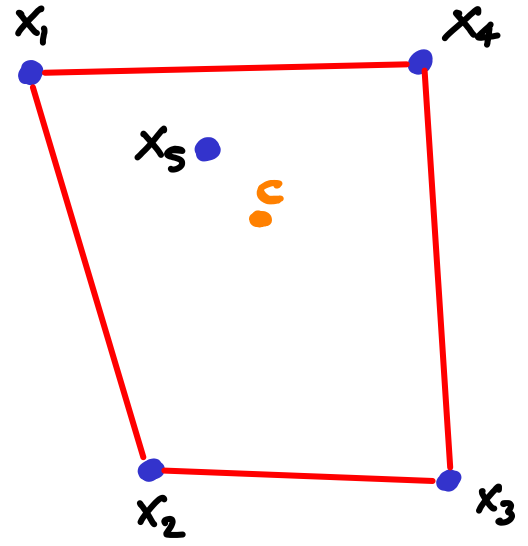
& $\sum_{i=1}^{n} \alpha_i = 1$

then $\cancel{\min\{x_i\}} \leq \sum_{i=1}^{n} \alpha_i x_i \leq \cancel{\max\{x_i\}}$

↳ is inside CONVEX HULL :

(a definition of "extreme")

ex: $\alpha_i = \frac{1}{n} \rightarrow \Sigma =$ center of mass, c



More about convexity & linear combinations in 2D

Given $x_1\ x_2\ x_3 \dots x_n \in \mathbb{R}^2$ (input)

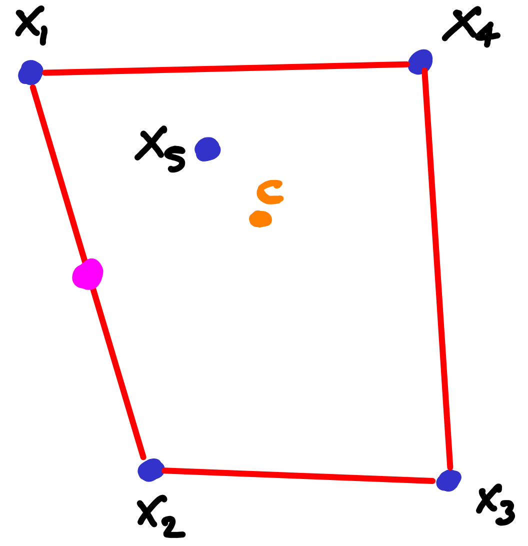
& $\alpha_1\ \alpha_2\ \alpha_3 \dots \alpha_n \in \mathbb{R}$ (weights) s.t $0 \leq \alpha_i \leq 1$

& $\sum_{i=1}^{n} \alpha_i = 1$

then $\cancel{\min\{x_i\}} \leq \sum_{i=1}^{n} \alpha_i x_i \leq \cancel{\max\{x_i\}}$

↳ is inside CONVEX HULL :

(a definition of "extreme")

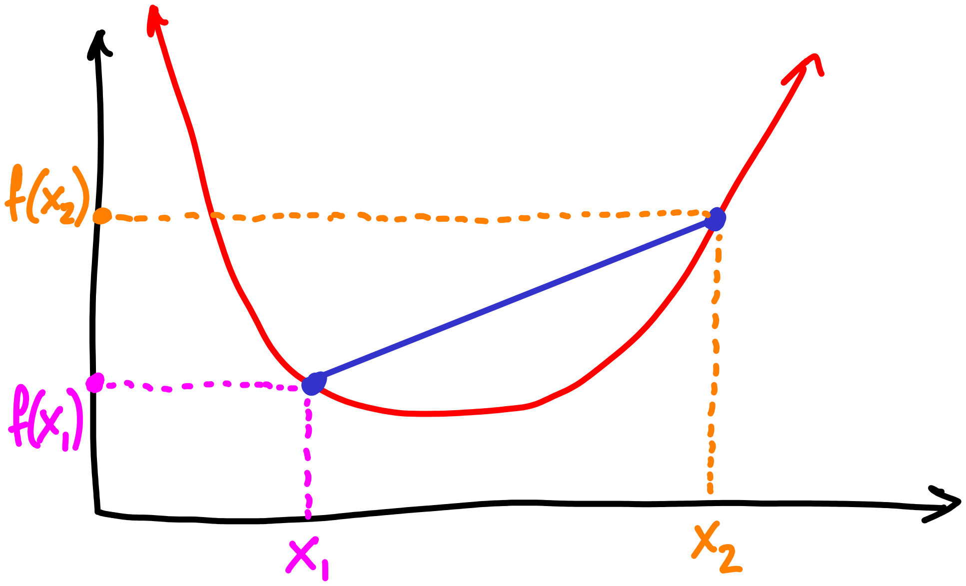


ex: $\alpha_i = \frac{1}{n} \rightarrow \Sigma =$ center of mass, c

ex2: $\alpha_1 = \alpha_2 = \frac{1}{2}$ =

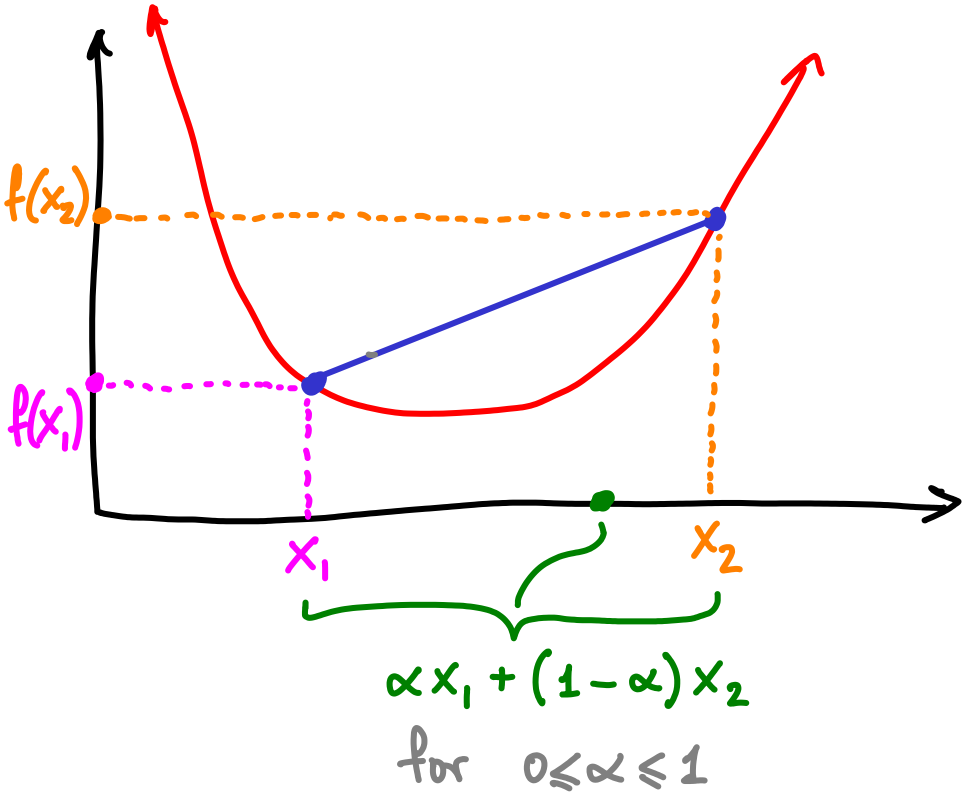
Convexity in terms of linear combinations

The segment joining any two points $[x_1, f(x_1)]$ & $[x_2, f(x_2)]$ must be ~~above~~ not below f.



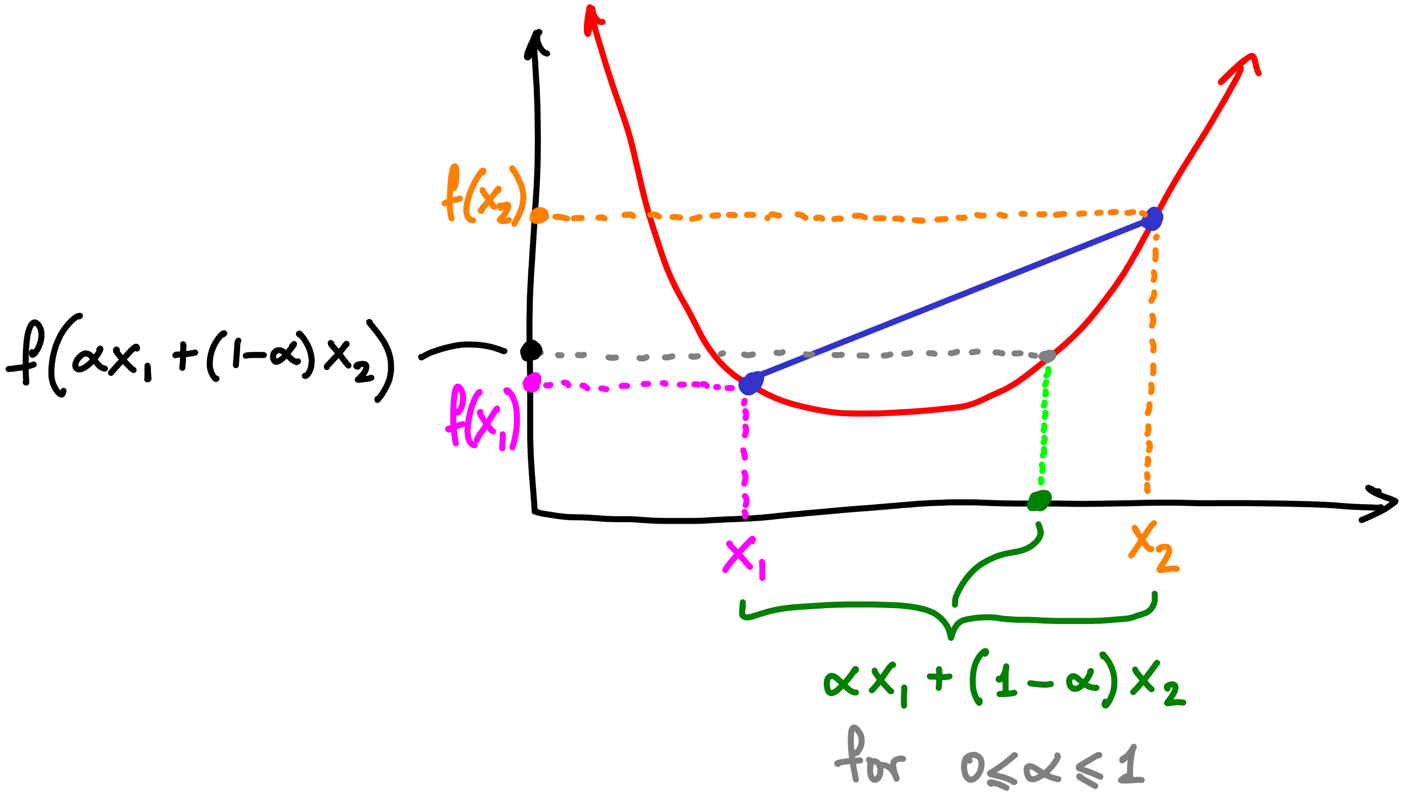
Convexity in terms of linear combinations

The segment joining any two points $[x_1, f(x_1)]$ & $[x_2, f(x_2)]$ must be above f.



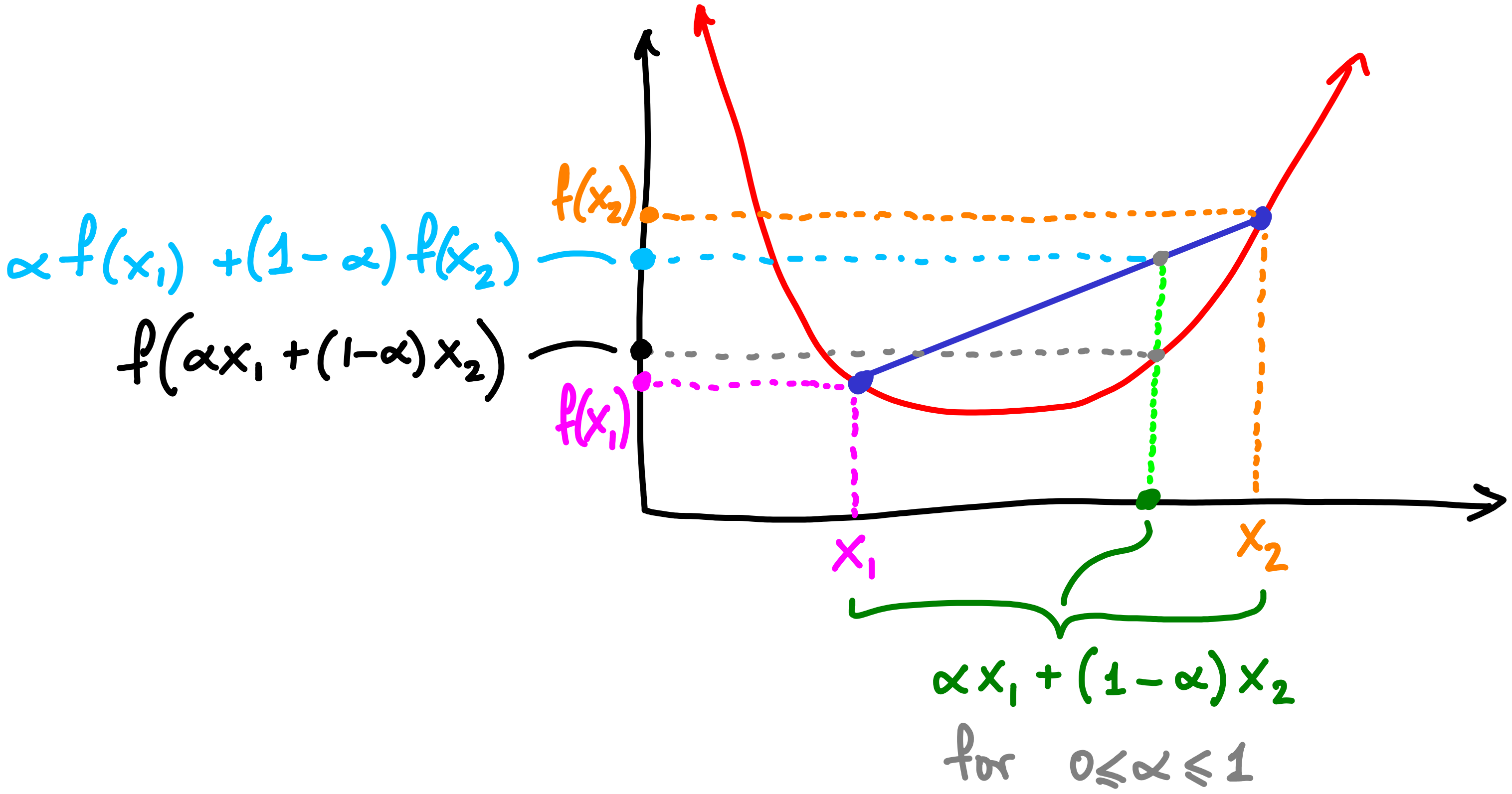
Convexity in terms of linear combinations

The segment joining any two points $[x_1, f(x_1)]$ & $[x_2, f(x_2)]$ must be above f.



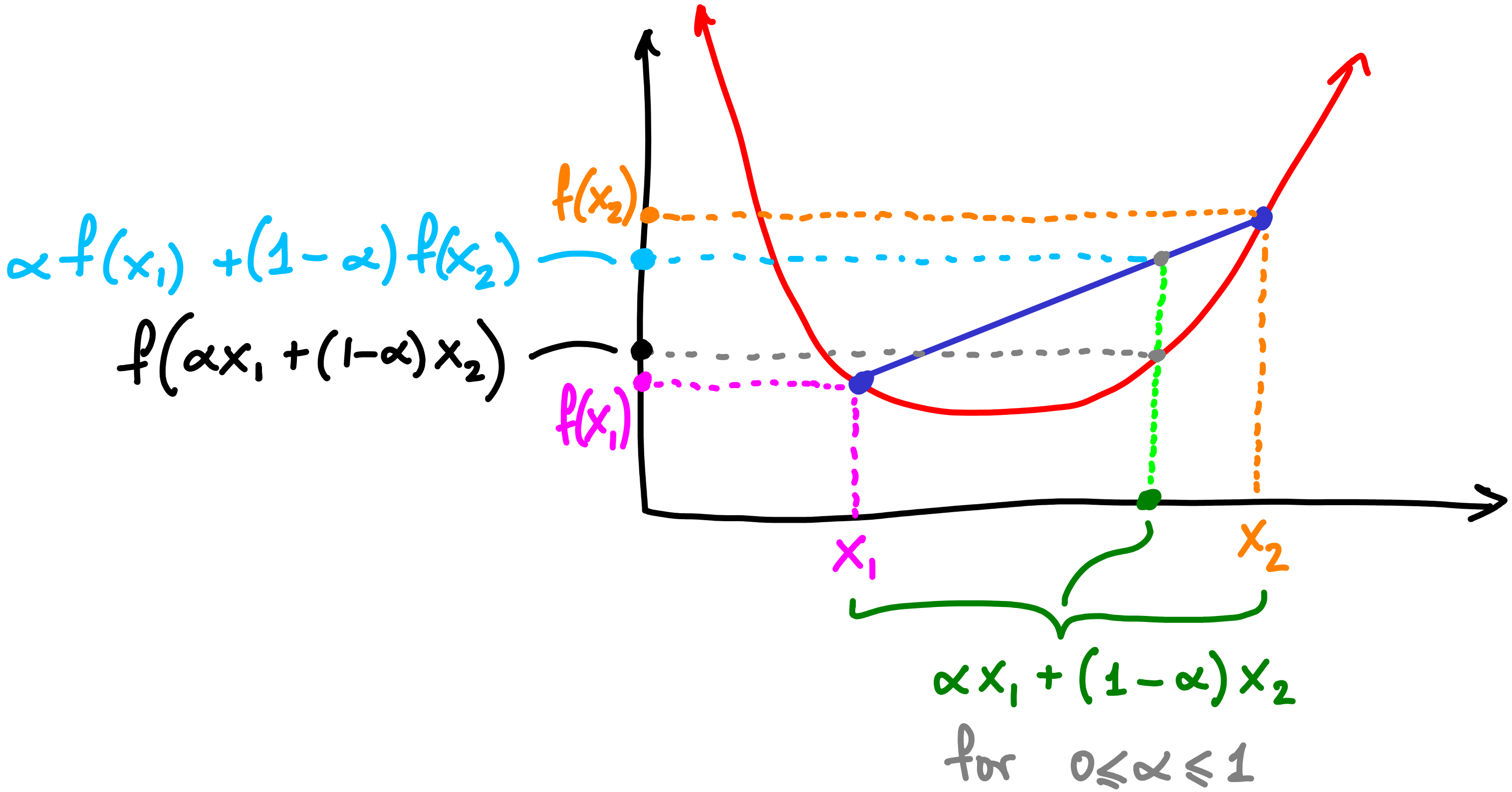
Convexity in terms of linear combinations

The segment joining any two points $[x_1, f(x_1)]$ & $[x_2, f(x_2)]$ must be above f.



Convexity in terms of linear combinations

The segment joining any two points $[x_1, f(x_1)]$ & $[x_2, f(x_2)]$ must be above f.



f is convex iff

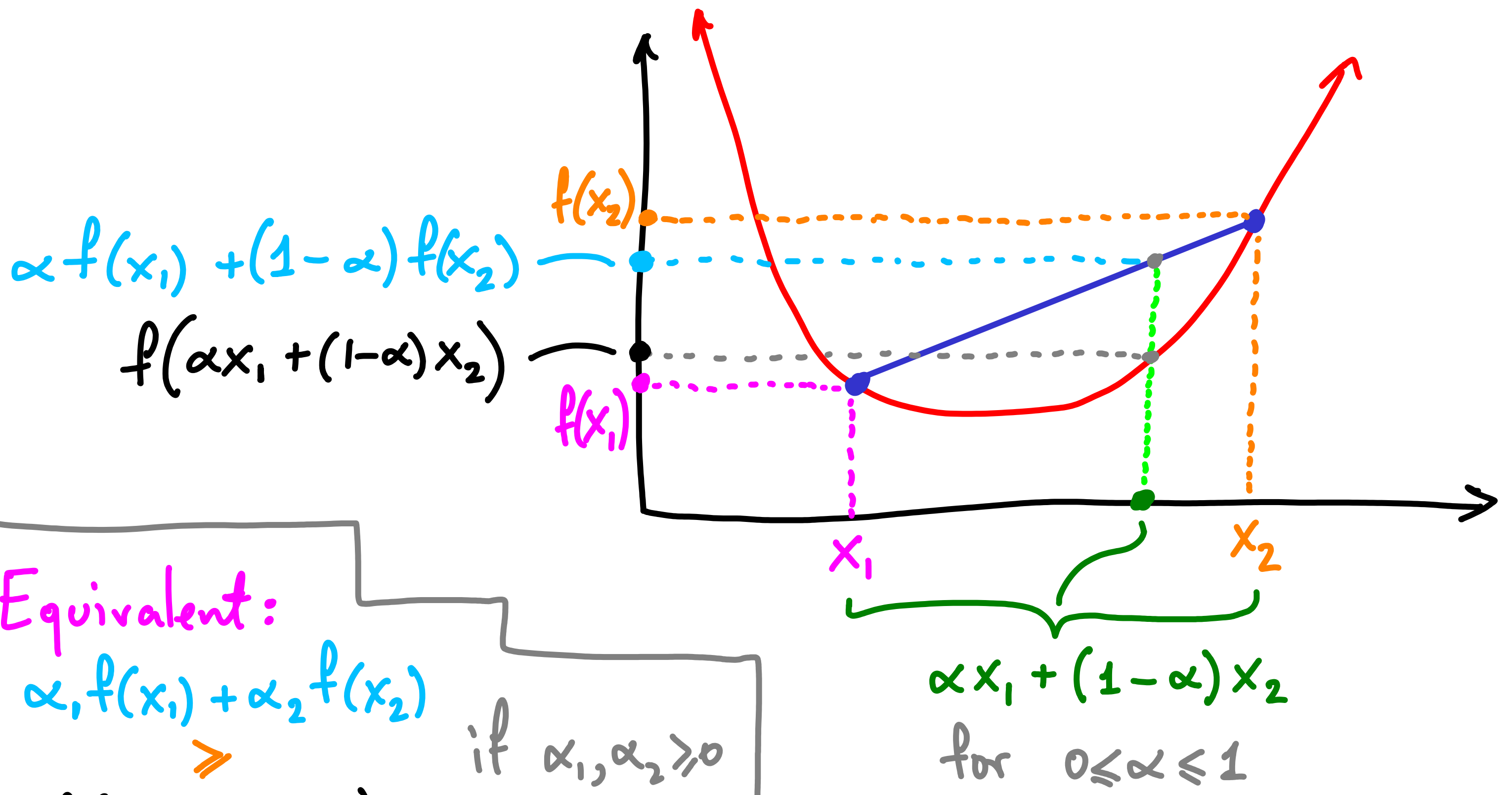
for all x_i, x_j

& all $0 \leq \alpha \leq 1$

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(\alpha x_1 + (1-\alpha) x_2)$$

Convexity in terms of linear combinations

The segment joining any two points $[x_1, f(x_1)]$ & $[x_2, f(x_2)]$ must be above f.



Equivalent:

$$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geqslant f(\alpha_1 x_1 + \alpha_2 x_2)$$

if $\alpha_1, \alpha_2 \geqslant 0$ & $\alpha_1 + \alpha_2 = 1$

f is convex iff for all x_i, x_j & all $0 \leqslant \alpha \leqslant 1$

$$\underbrace{\alpha f(x_1) + (1-\alpha) f(x_2) \geqslant f(\alpha x_1 + (1-\alpha) x_2)}$$

(2D definition)

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geqslant f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geqslant 0$ & $\alpha_1 + \alpha_2 = 1$ } def. convex f.

we are heading to

$$f(E[x]) \leqslant E[f(x)]$$

(Jensen)

$$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geqslant f(\alpha_1 x_1 + \alpha_2 x_2) \quad \text{if } \alpha_1, \alpha_2 \geqslant 0 \;\&\; \alpha_1 + \alpha_2 = 1 \Big\} \text{ def. convex f.}$$

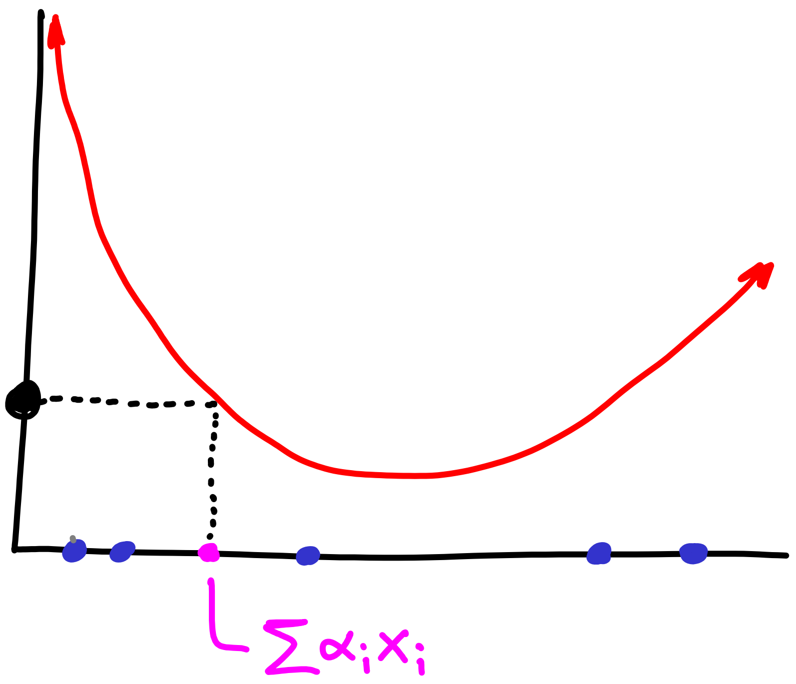
Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leqslant \sum_{i=1}^{n} \alpha_i f(x_i)$$

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geqslant f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geqslant 0$ & $\alpha_1 + \alpha_2 = 1$ $\Big\}$ def. convex f.

Generalization:

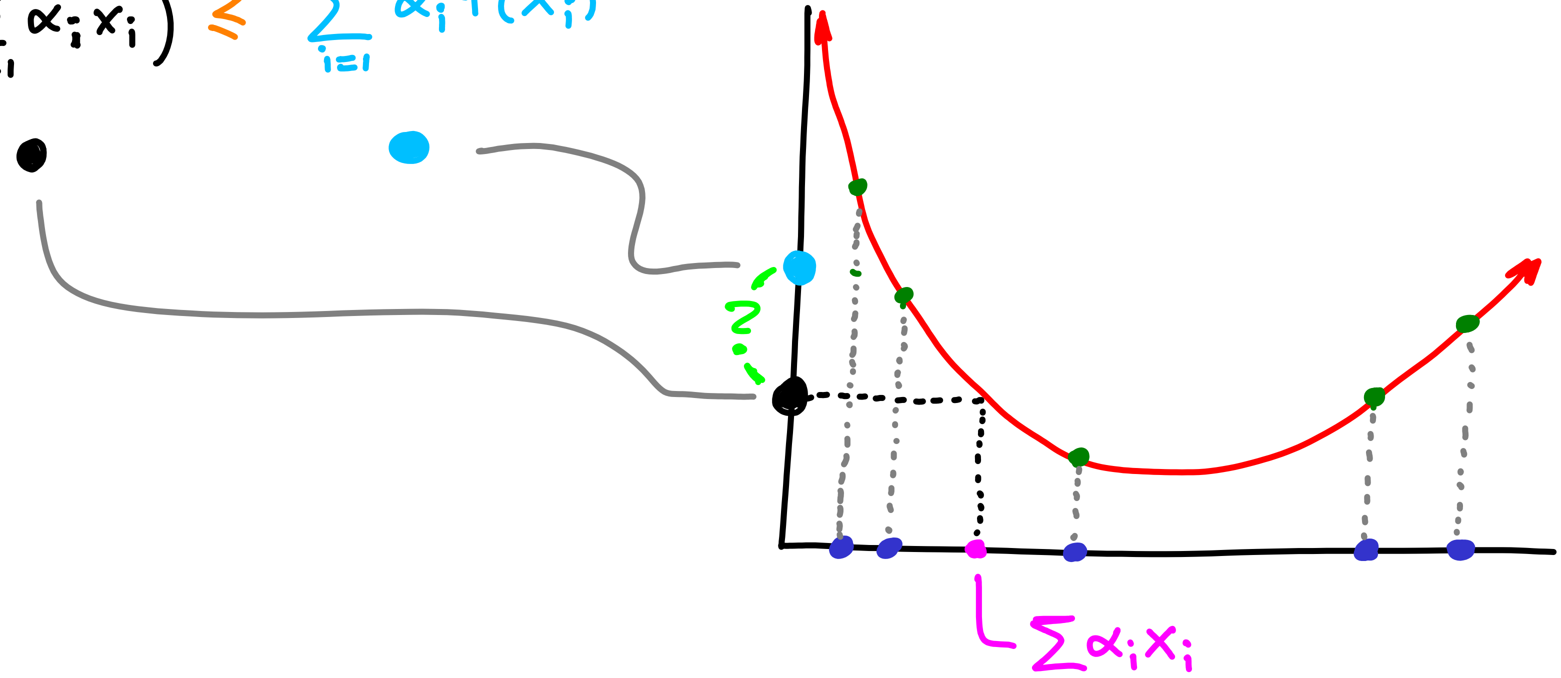
$$\underbrace{f\Big(\underbrace{\sum_{i=1}^{n} \alpha_i x_i}\Big)} \leqslant \sum_{i=1}^{n} \alpha_i f(x_i)$$



$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geq f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geq 0$ & $\alpha_1 + \alpha_2 = 1$ $\Big\}$ def. convex f.

Generalization:

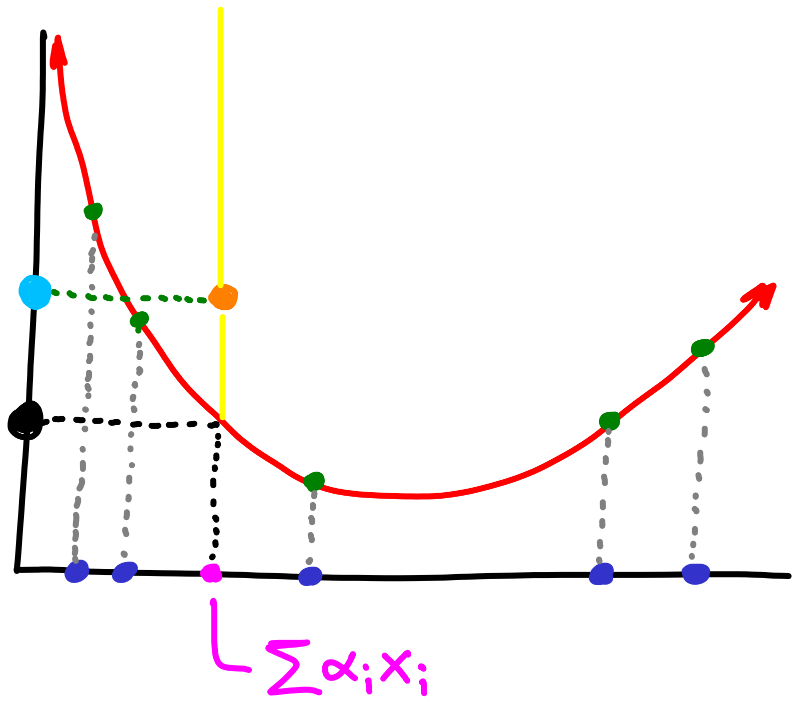
$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$



$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geqslant f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geqslant 0$ & $\alpha_1 + \alpha_2 = 1$ } def. convex f.

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leqslant \sum_{i=1}^{n} \alpha_i f(x_i)$$



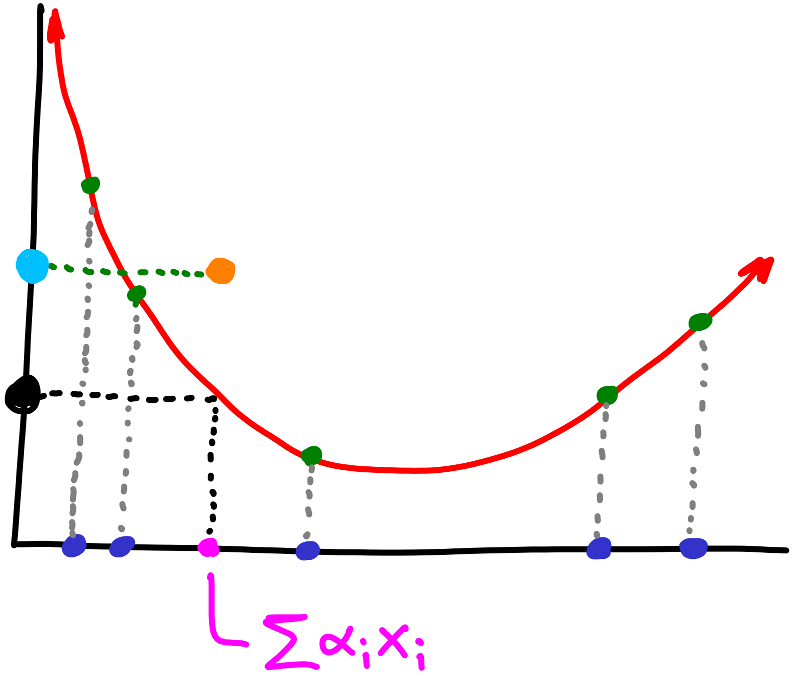
● = $\underbrace{\sum \alpha_i x_i}_{\text{on same vertical line as } \sum \alpha_i x_i}, \sum \alpha_i f(x_i)$

(still don't know if above ●)

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geqslant f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geqslant 0$ & $\alpha_1 + \alpha_2 = 1$ } def. convex f.

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leqslant \sum_{i=1}^{n} \alpha_i f(x_i)$$



● $= \underbrace{\sum \alpha_i x_i}$, $\sum \alpha_i f(x_i)$

on same vertical line as $\sum \alpha_i x_i$

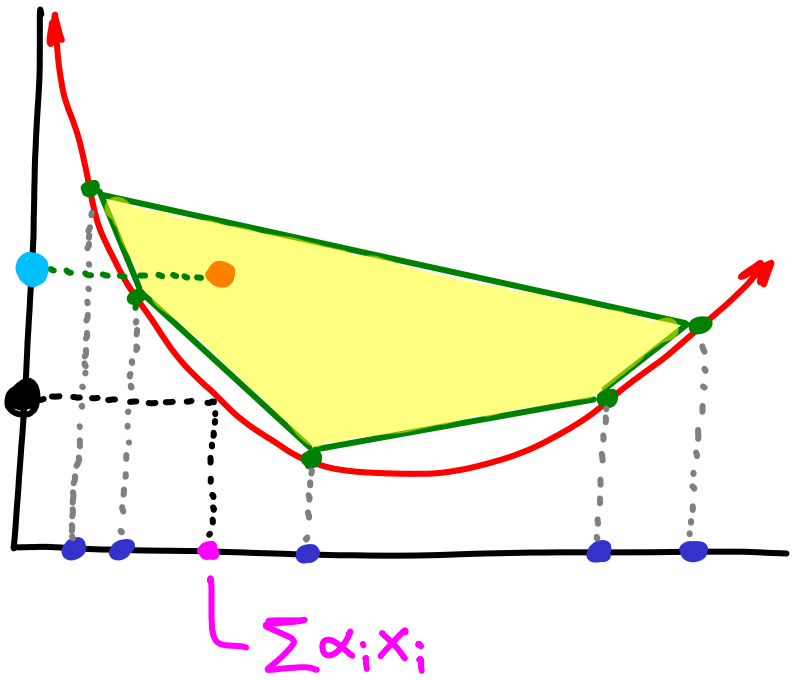
(still don't know if above ●)

$\approx \sum \alpha_i (x_i, f(x_i))$

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geq f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geq 0$ & $\alpha_1 + \alpha_2 = 1$ } def. convex f.

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$



● = $\sum \alpha_i x_i$, $\sum \alpha_i f(x_i)$

on same vertical line as $\sum \alpha_i x_i$

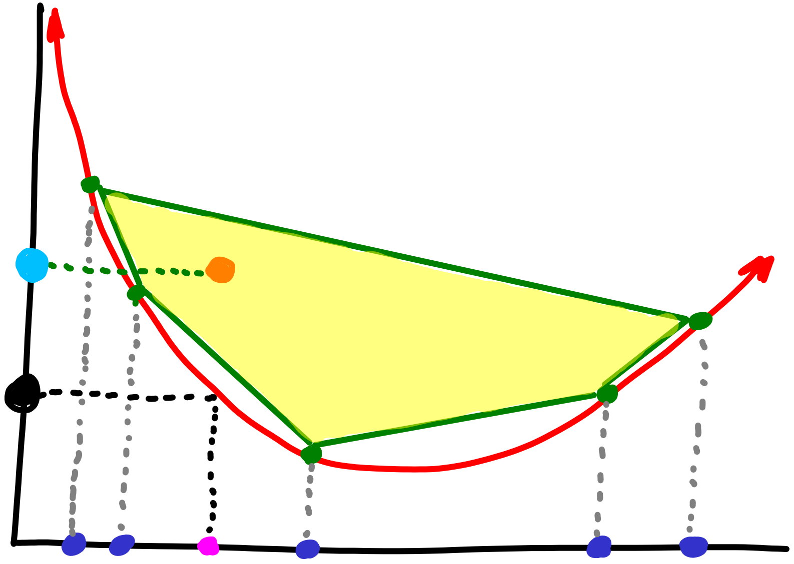
(still don't know if above ●)

≈ $\sum \alpha_i (x_i, f(x_i))$ → inside convex hull of all $(x_i, f(x_i))$

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geq f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geq 0$ & $\alpha_1 + \alpha_2 = 1$ } def. convex f.

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$



$\Sigma \alpha_i x_i$

$\bullet = \left(\underbrace{\Sigma \alpha_i x_i}, \Sigma \alpha_i f(x_i)\right)$

on same vertical line as $\Sigma \alpha_i x_i$

(still don't know if above $\bullet$)

$\approx \Sigma \alpha_i (x_i, f(x_i)) \longrightarrow$ inside convex hull of all $(x_i, f(x_i))$ }

$\alpha_1 f(x_1) + \alpha_2 f(x_2) \geq f(\alpha_1 x_1 + \alpha_2 x_2)$ if $\alpha_1, \alpha_2 \geq 0$ & $\alpha_1 + \alpha_2 = 1$ } def. convex f.

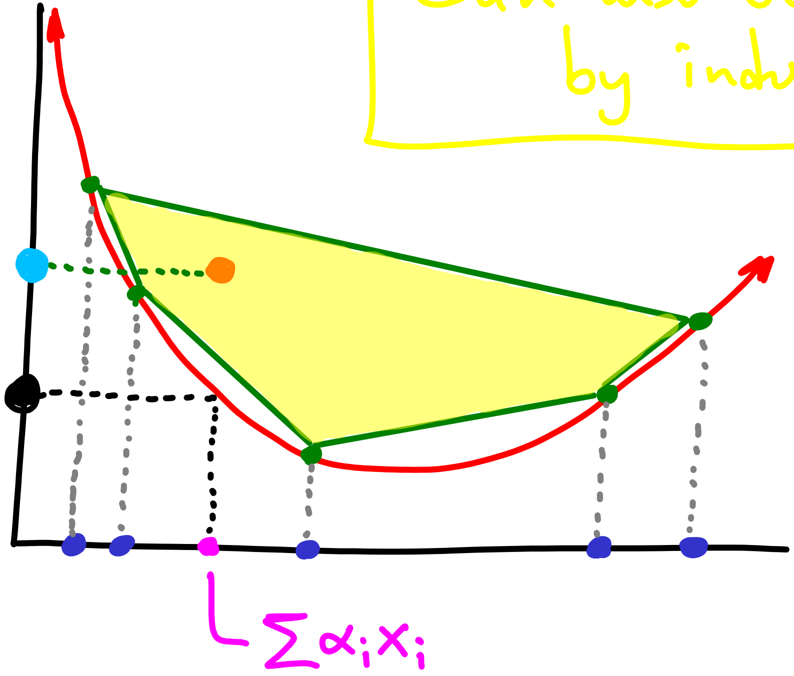
Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$

$\bullet \leq \bullet$

QED

Can also be proved by induction



$\bullet = \left(\underbrace{\sum \alpha_i x_i}, \sum \alpha_i f(x_i)\right)$

on same vertical line as $\sum \alpha_i x_i$

(still don't know if above $\bullet$)

$= \sum \alpha_i (x_i, f(x_i)) \longrightarrow$ inside convex hull of all $(x_i, f(x_i))$ } $\bullet$ } $\bullet \geq \bullet$

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) = f\left(\alpha_n x_n + \underline{(1-\alpha_n)} \sum_{i=1}^{n-1} \frac{\alpha_i}{\underline{1-\alpha_n}} x_i\right)$$

$$\sum_{i=1}^{n-1} \frac{\alpha_i}{1-\alpha_n} = 1$$

keep it normalized so that inductive hypothesis can be applied

Generalization:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) = f\left(\alpha_n x_n + (1-\alpha_n)\sum_{i=1}^{n-1} \frac{\alpha_i}{1-\alpha_n} x_i\right)$$

$$\sum_{i=1}^{n-1} \frac{\alpha_i}{1-\alpha_n} = 1$$

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\cdot f\left(\sum_{i=1}^{n-1} \frac{\alpha_i}{1-\alpha_n} x_i\right)$$

by induction, if base: $n=2$

...

Generalization:

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) \leq \sum_{i=1}^{n}\alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

$$\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n} = 1$$

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) = f\Big(\underbrace{\alpha_n x_n}_{} + \underbrace{(1-\alpha_n)\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i}_{}\Big)$$

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\cdot f\left(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

by induction, if base: $n=2$

$$\left.\begin{aligned}\alpha_n &= \alpha'_1 \geq 0\\ 1-\alpha_n &= \alpha'_2 \geq 0\end{aligned}\right\} \alpha'_1+\alpha'_2 = 1$$

$$f(\alpha'_1 \cdot x'_1 + \alpha'_2 \cdot x'_2)$$

$$\leq \alpha'_1 f(x'_1) + \alpha'_2 f(x'_2)$$

$$x_n = x'_1 \qquad \sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i = x'_2$$

Generalization:

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) \leq \sum_{i=1}^{n}\alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) = f\left(\alpha_n x_n + (1-\alpha_n)\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

$$\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n} = 1$$

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\cdot f\left(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

by convexity

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\cdot\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}f(x_i)$$

by induction

Generalization:

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) \leq \sum_{i=1}^{n}\alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) = f\left(\alpha_n x_n + (1-\alpha_n)\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

because $\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n} = 1$

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\cdot f\left(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

by convexity

$$\leq \alpha_n f(x_n) + \underline{(1-\alpha_n)}\cdot\sum_{i=1}^{n-1}\frac{\alpha_i}{\underline{1-\alpha_n}}f(x_i)$$

by induction

$$= \alpha_n f(x_n) + \sum_{i=1}^{n-1}\alpha_i\cdot f(x_i)$$

temporary use of $(1-\alpha_n)$

Generalization:

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) \leq \sum_{i=1}^{n}\alpha_i f(x_i)$$

(for convex f)

(Another) Proof by induction.

We've seen it's true for $n=2$ and a trivial base case is $n=1$

$$f\left(\sum_{i=1}^{n}\alpha_i x_i\right) = f\left(\alpha_n x_n + (1-\alpha_n)\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

because $\displaystyle\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n} = 1$

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\cdot f\left(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i\right)$$

by convexity

$$\leq \alpha_n f(x_n) + (1-\alpha_n)\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}f(x_i)$$

by induction

$$= \alpha_n f(x_n) + \sum_{i=1}^{n-1}\alpha_i\cdot f(x_i)$$

$$= \sum_{i=1}^{n}\alpha_i f(x_i)$$

QED

$$\boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)}$$

we just proved this, for convex f

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)$$

Finally, Jensen's inequality, for convex f:

$$f(E[x]) \leq E[f(x)]$$

(discrete version)

$$\boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)}$$

Finally, Jensen's inequality, for convex f:

$$f(E[x]) \leq E[f(x)]$$ (discrete version)

$$f(E[x]) = f\Big(\underbrace{\sum_{j=-\infty}^{\infty} j \cdot P[x=j]}_{\text{def.}}\Big)$$

$$\boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)} \quad (1)$$

Finally, Jensen's inequality, for convex f:

$$f(E[x]) \leq E[f(x)]$$ (discrete version)

$$f(E[x]) = f\left(\sum_{j=-\infty}^{\infty} j \cdot P[x=j]\right)$$

$$\leq \sum_{j=-\infty}^{\infty} P[x=j] \cdot f(j)$$ by (1) (probabilities sum to 1 $\rightarrow$ like α_i)

(1) $\boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)}$

Finally, Jensen's inequality, for convex f:

$$f(E[x]) \leq E[f(x)]$$

(discrete version)

$$f(E[x]) = f\left(\sum_{j=-\infty}^{\infty} j \cdot P[x=j]\right)$$

$$\leq \sum_{j=-\infty}^{\infty} P[x=j] \cdot f(j)$$

by (1) (probabilities sum to 1 → like α_i)

$$\sum_{y} y \cdot P[f(x)=y] = E[f(x)]$$

y: all possible values of $f(x)$

$$\text{(1)} \quad \boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)}$$

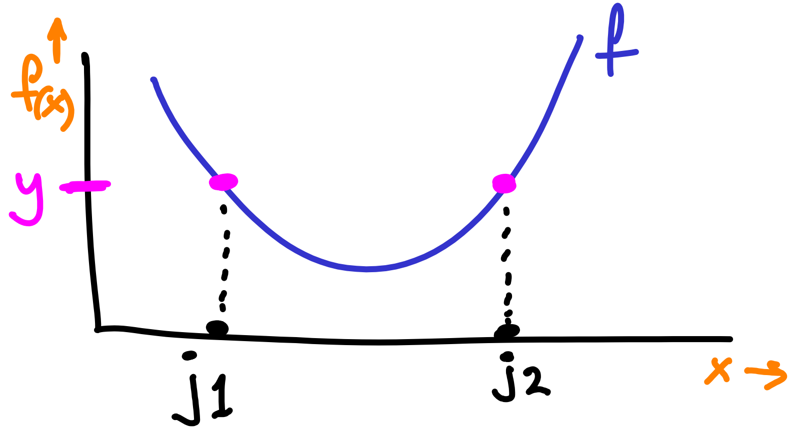
Finally, Jensen's inequality, for convex f:

$$f(E[x]) \leq E[f(x)] \quad \text{(discrete version)}$$

$$f(E[x]) = f\left(\sum_{j=-\infty}^{\infty} j \cdot P[x=j]\right)$$

$$\leq \sum_{j=-\infty}^{\infty} P[x=j] \cdot f(j) \quad \text{by (1) (probabilities sum to 1} \rightarrow \text{like } \alpha_i)$$

$$\underbrace{\sum_{\substack{\text{all } j \\ \text{s.t. } f(j)=y}} P[x=j]}_{P[f(x)=y]} \qquad \sum_{y} y \cdot \underbrace{P[f(x)=y]} = E[f(x)]$$



y: all possible values of $f(x)$

(1) $$\boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \le \sum_{i=1}^{n} \alpha_i f(x_i)}$$

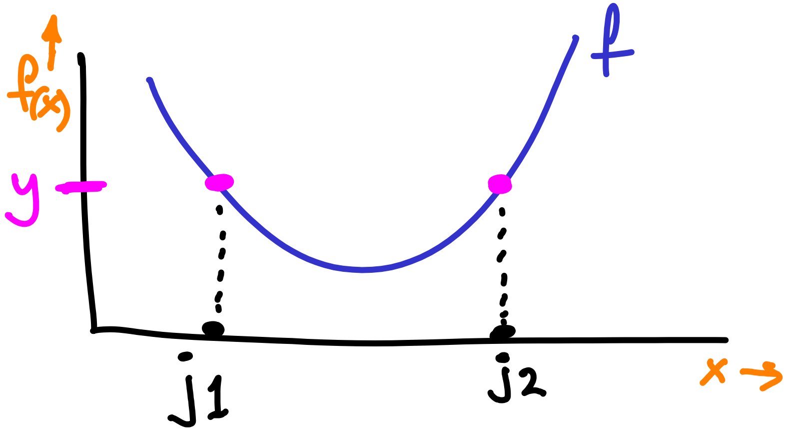
Finally, Jensen's inequality, for convex f:

$$f(E[x]) \le E[f(x)]$$ (discrete version)

$$f(E[x]) = f\left(\sum_{j=-\infty}^{\infty} j \cdot P[x=j]\right)$$

$$\le \sum_{j=-\infty}^{\infty} P[x=j] \cdot f(j)$$ by (1) (probabilities sum to 1 $\to$ like α_i)

$$\sum_{\text{all } y}\left(y \cdot \underbrace{\sum_{\substack{\text{all } j \\ \text{s.t. } f(j)=y}} P[x=j]}_{P[f(x)=y]}\right) = \sum_{y} y \cdot \underbrace{P[f(x)=y]} = E[f(x)]$$



y: all possible values of $f(x)$

$$(1)\quad \boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i)}$$

Finally, Jensen's inequality, for convex f:

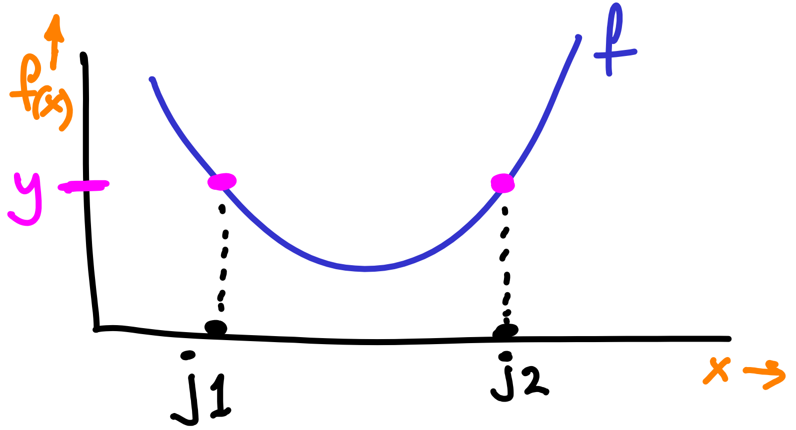
$$f(E[x]) \leq E[f(x)] \quad \text{(discrete version)}$$

$$f(E[x]) = f\left(\sum_{j=-\infty}^{\infty} j \cdot P[x=j]\right)$$

$$\leq \sum_{j=-\infty}^{\infty} P[x=j] \cdot f(j) \quad \text{by (1) (probabilities sum to 1} \rightarrow \text{like } \alpha_i)$$

$$= \sum_{\text{all } y}\Bigg(y \cdot \underbrace{\sum_{\substack{\text{all } j \\ \text{s.t. } f(j)=y}} P[x=j]}_{P[f(x)=y]}\Bigg) = \sum_{y} y \cdot \underbrace{P[f(x)=y]}_{} = E[f(x)]$$

y: all possible values of $f(x)$



$E[\text{height of random BST}] = O(\log n)$ PROOF (PART 2)

$E[\text{height of random BST}] = O(\log n)$ PROOF (PART 2)

DONE (1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

$E[\text{height of random BST}] = O(\log n)$ PROOF (PART 2)

DONE (1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

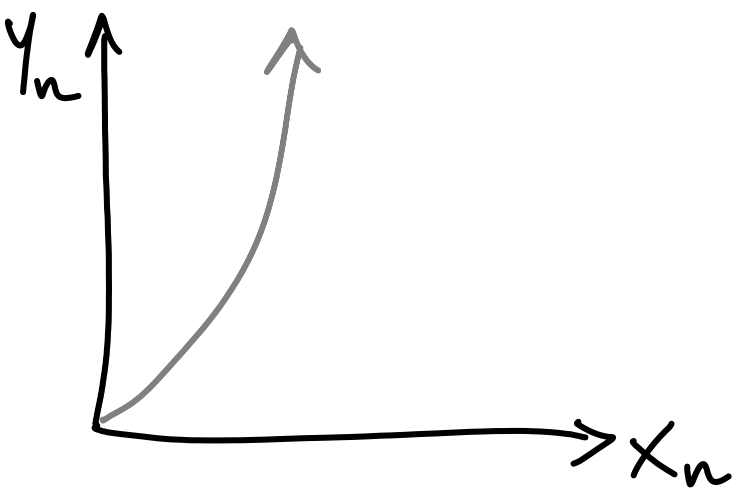
NOW: (2) For X_n = random variable = height of random tree w/ n nodes,
prove $E[2^{X_n}] = O(n^3)$

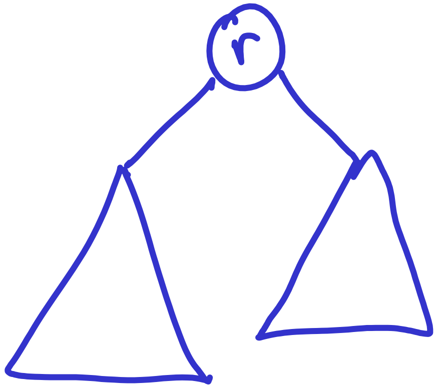
$E[\text{height of random BST}] = O(\log n)$ PROOF (PART 2)

DONE (1) Jensen's inequality for convex f: $f(E[x]) \leq E[f(x)]$

NOW: (2) For X_n = random variable = height of random tree w/ n nodes, prove $E[2^{X_n}] = O(n^3)$ Let $2^{X_n} = Y_n$

this is convex.

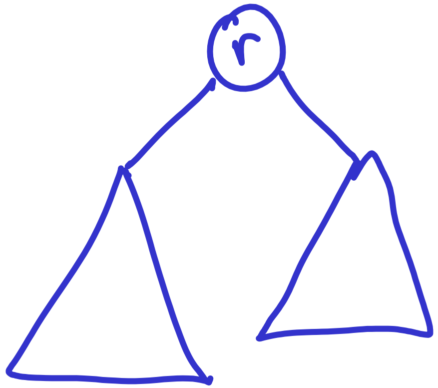




After placing a root, recurse twice:

2 random BST's

↪ contents depend on r.

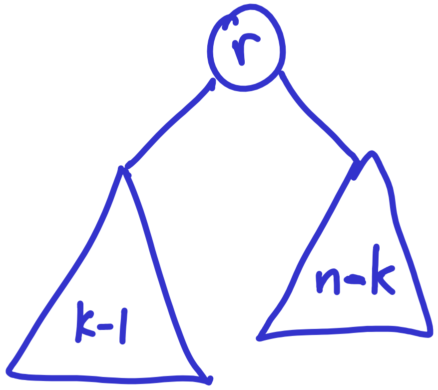


After placing a root, recurse twice:

2 random BST's

↳ contents depend on r.

↳ structure depends on size.



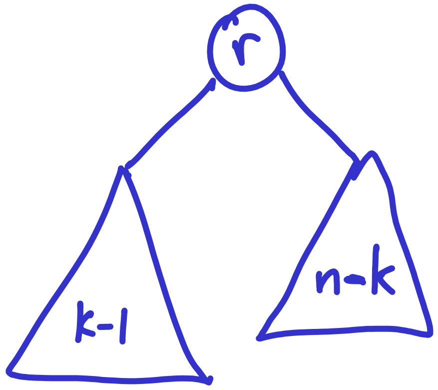
After placing a root, recurse twice:

2 random BST's

↳ contents depend on r.

↳ structure depends on size.

Let r have rank k.



After placing a root, recurse twice:

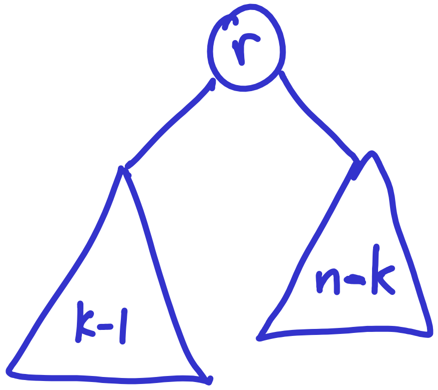
2 random BST's

↳ contents depend on r.

↳ structure depends on size.

Let r have rank k.

↳ $X_n = 1 + \max\{X_{k-1}, X_{n-k}\}$



After placing a root, recurse twice:

2 random BST's

↳ contents depend on r.

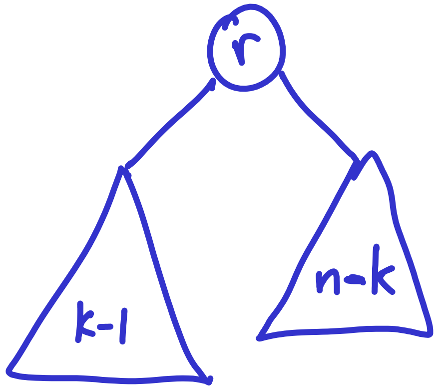
↳ structure depends on size.

Let r have rank k.

↳ $X_n = 1 + \max\{X_{k-1}, X_{n-k}\}$

$Y_n = 2^{1+\max\{\cdot\}}$

$$Y_n = 2^{1+\max\{X_{k-1}, X_{n-k}\}}$$



After placing a root, recurse twice:

2 random BST's

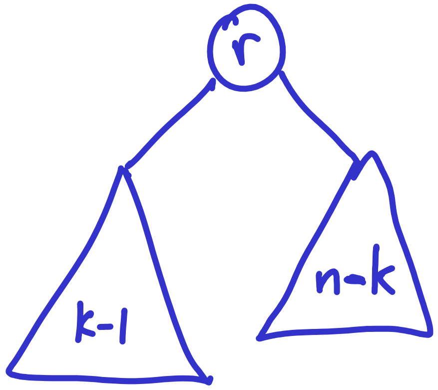
↳ contents depend on r.

↳ structure depends on size.

Let r have rank k.

↳ $X_n = 1 + \max\{X_{k-1}, X_{n-k}\}$

$Y_n = 2^{1+\max\{\cdot\}} = 2 \cdot 2^{\max\{\cdot\}}$



After placing a root, recurse twice:

2 random BST's

↳ contents depend on r.

↳ structure depends on size.

Let r have rank k.

↳ $X_n = 1 + \max\{X_{k-1}, X_{n-k}\}$

$$Y_n = 2^{1+\max\{\cdot\}} = 2 \cdot 2^{\max\{\cdot\}} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$$

$$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$$ (for fixed root w/ rank k)

$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$ (for fixed root w/ rank k)

Define $Z_{nk} = \begin{cases} 1 & \text{if root has rank } k \text{ of } n \\ 0 & \text{otherwise} \end{cases}$

$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$ (for fixed root w/ rank k)

Define $Z_{nk} = \begin{cases} 1 & \text{if root has rank } k \text{ of } n \\ 0 & \text{otherwise} \end{cases}$

$E[Z_{nk}] = 1 \cdot P\{Z_{nk}=1\} + 0 \cdot P\{Z_{nk}=0\}$

$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$ (for fixed root w/ rank k)

Define $Z_{nk} = \begin{cases} 1 & \text{if root has rank } k \text{ of } n \\ 0 & \text{otherwise} \end{cases}$

$E[Z_{nk}] = 1 \cdot P\{Z_{nk}=1\} + 0 \cdot P\{Z_{nk}=0\} = P\{Z_{nk}=1\} = \frac{1}{n}$

$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$ (for fixed root w/ rank k)

Define $Z_{nk} = \begin{cases} 1 & \text{if root has rank } k \text{ of } n \\ 0 & \text{otherwise} \end{cases}$

$E[Z_{nk}] = 1 \cdot P\{Z_{nk}=1\} + 0 \cdot P\{Z_{nk}=0\} = P\{Z_{nk}=1\} = \frac{1}{n}$

$Y_n = \sum_{k=1}^{n} \left(Z_{nk} \cdot 2\max\{Y_{k-1}, Y_{n-k}\} \right)$

unspecified root

$Y_n = \text{one of} \begin{cases} k=1 \rightarrow 2 \cdot \max\{Y_0, Y_{n-1}\} \\ k=2 \rightarrow 2 \cdot \max\{Y_1, Y_{n-2}\} \\ \vdots \\ \vdots \\ k=n \rightarrow 2 \cdot \max\{Y_{n-1}, Y_0\} \end{cases}$

$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$ (for fixed root w/ rank k)

Define $Z_{nk} = \begin{cases} 1 & \text{if root has rank } k \text{ of } n \\ 0 & \text{otherwise} \end{cases}$

$E[Z_{nk}] = 1 \cdot P\{Z_{nk}=1\} + 0 \cdot P\{Z_{nk}=0\} = P\{Z_{nk}=1\} = \frac{1}{n}$

$Y_n = \sum_{k=1}^{n} \left(Z_{nk} \cdot 2\max\{Y_{k-1}, Y_{n-k}\} \right)$

$E[Y_n] = \sum_{k=1}^{n} E\left[Z_{nk} \cdot 2\max\{Y_{k-1}, Y_{n-k}\} \right]$ by linearity of expectation

$Y_n = 2^{X_n} = 2 \cdot \max\{Y_{k-1}, Y_{n-k}\}$ (for fixed root w/ rank k)

Define $Z_{nk} = \begin{cases} 1 & \text{if root has rank } k \text{ of } n \\ 0 & \text{otherwise} \end{cases}$

$E[Z_{nk}] = 1 \cdot P\{Z_{nk}=1\} + 0 \cdot P\{Z_{nk}=0\} = P\{Z_{nk}=1\} = \frac{1}{n}$

$$Y_n = \sum_{k=1}^{n} \left(Z_{nk} \cdot 2\max\{Y_{k-1}, Y_{n-k}\} \right)$$

$E[Y_n] = \sum_{k=1}^{n} E\left[Z_{nk} \cdot 2\max\{Y_{k-1}, Y_{n-k}\}\right]$ by linearity of expectation

$= 2 \sum E[Z_{nk}] \cdot E\left[\max\{Y_{k-1}, Y_{n-k}\}\right]$ by independence

$$E[Y_n] = 2\sum_{k=1}^{n} E[Z_{nk}] \cdot E\left[\max\{Y_{k-1}, Y_{n-k}\}\right]$$

$$E[Y_n] = 2 \sum_{k=1}^{n} E[Z_{nk}] \cdot E\left[\max\{Y_{k-1}, Y_{n-k}\}\right]$$

$$\le 2 \cdot \frac{1}{n} \cdot \sum_{k=1}^{n} E\left[Y_{k-1} + Y_{n-k}\right]$$

because $\max(a,b) \le a+b$

$$E[Y_n] = 2\sum_{k=1}^{n} E[Z_{nk}] \cdot E\left[\max\{Y_{k-1}, Y_{n-k}\}\right]$$

$$\leq 2 \cdot \frac{1}{n} \cdot \sum_{k=1}^{n} E[Y_{k-1} + Y_{n-k}]$$ because $\max(a,b) \leq a+b$

$$= \frac{2}{n}\sum_{k=1}^{n} E[Y_{k-1}] + E[Y_{n-k}]$$ ...linearity exp.

$$E[Y_n] = 2\sum_{k=1}^{n} E[Z_{nk}] \cdot E\left[\max\{Y_{k-1}, Y_{n-k}\}\right]$$

$$\leq 2 \cdot \frac{1}{n} \cdot \sum_{k=1}^{n} E[Y_{k-1} + Y_{n-k}]$$ because $\max(a,b) \leq a+b$

$$= \frac{2}{n}\sum_{k=1}^{n} E[Y_{k-1}] + E[Y_{n-k}]$$ ...linearity exp.

$$\leq \frac{4}{n}\sum_{k=0}^{n-1} E[Y_k]$$ by symmetry (also notice $\sum_{k=0}^{n-1}$ change)

$$E[Y_n] = 2\sum_{k=1}^{n} E[Z_{nk}]\cdot E\left[\max\{Y_{k-1}, Y_{n-k}\}\right]$$

$$\leq 2\cdot\frac{1}{n}\cdot\sum_{k=1}^{n} E[Y_{k-1} + Y_{n-k}]$$ because $\max(a,b) \leq a+b$

$$= \frac{2}{n}\sum_{k=1}^{n} E[Y_{k-1}] + E[Y_{n-k}]$$ ...linearity exp.

$$\leq \frac{4}{n}\sum_{k=0}^{n-1} E[Y_k]$$ by symmetry (also notice $\sum_{k=0}^{n-1}$ change)

Claim $E[Y_n] \leq c\cdot n^3$

Base case $n=1$: $E[Y_1] = \text{const.}$

$E[2^{X_1}] = 2 < c\cdot 1^3$

ok

(X_1 = height of tree of size 1)

$$E[Y_n] \leq \frac{4}{n} \sum_{k=0}^{n-1} E[Y_k]$$

Claim $E[Y_n] \leq c \cdot n^3$

$$E[Y_n] \leq \frac{4}{n} \sum_{k=0}^{n-1} E[Y_k]$$

Claim $E[Y_n] \leq c \cdot n^3$

$$\leq \frac{4}{n} \sum_{k=0}^{n-1} c \cdot k^3$$

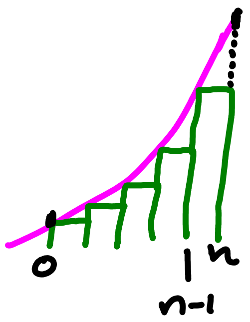
hypothesis

$$E[Y_n] \leq \frac{4}{n} \sum_{k=0}^{n-1} E[Y_k]$$

Claim $E[Y_n] \leq c \cdot n^3$

$$\leq \frac{4}{n} \sum_{k=0}^{n-1} c \cdot k^3$$

$$\leq \frac{4c}{n} \int_0^n k^3 \, dk$$

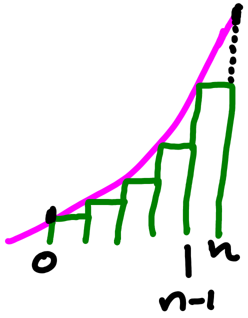


$$E[Y_n] \leq \frac{4}{n} \sum_{k=0}^{n-1} E[Y_k]$$

Claim $E[Y_n] \leq c \cdot n^3$: true

$$\leq \frac{4}{n} \sum_{k=0}^{n-1} c \cdot k^3$$

$$\leq \frac{4c}{n} \int_0^n k^3 \, dk$$



$$= \frac{4c}{n} \cdot \frac{n^4}{4} = cn^3$$

QED

$$\sum_{k=0}^{n-1} k^3 \leq \frac{n^4}{4}$$

Assume true for n-2:

$$(n-1)^3 + \sum_{k=0}^{n-2} k^3 \leq (n-1)^3 + \frac{(n-1)^4}{4}$$

$$= \frac{n^4 - 6n^2 + 8n - 3}{4} \quad \square$$

$$E[Y_n] \leq c \cdot n^3 = O(n^3)$$

$E[Y_n] \leq c \cdot n^3 = O(n^3)$

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$E[Y_n] \leq c \cdot n^3 = O(n^3)$

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$2^{E[X_n]} \leq E[2^{X_n}]$

$$E[Y_n] \leq c \cdot n^3 = O(n^3)$$

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$$2^{E[X_n]} \leq E[2^{X_n}] = E[Y_n] \leq c \cdot n^3$$

$$E[Y_n] \leq c \cdot n^3 = O(n^3)$$

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$$2^{E[X_n]} \leq E[2^{X_n}] = E[Y_n] \leq c \cdot n^3$$

$$\log 2^{E[X_n]} \leq \log(cn^3)$$

$E[Y_n] \leq c \cdot n^3 = O(n^3)$

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$2^{E[X_n]} \leq E[2^{X_n}] = E[Y_n] \leq c \cdot n^3$

$\log 2^{E[X_n]} \leq \log(cn^3)$

$E[X_n] \leq \log c + \log n^3$

$E[Y_n] \leq c \cdot n^3 = O(n^3)$ $\longrightarrow$ notice any $O(n^k)$ would give $k \log n$ below.

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$$2^{E[X_n]} \leq E[2^{X_n}] = E[Y_n] \leq c \cdot n^3$$

$$\log 2^{E[X_n]} \leq \log(cn^3)$$

$$\underline{\underline{E[X_n]}} \leq \log c + \log n^3$$

$$\leq \underline{\underline{3 \log n + O(1)}}$$

we even get the leading constant!

$E[Y_n] \leq c \cdot n^3 = O(n^3)$ $\longrightarrow$ notice any $O(n^k)$ would give $k \log n$ below.

Recall $f(E[x]) \leq E[f(x)]$ Jensen's inequality.

$$2^{E[X_n]} \leq E[2^{X_n}] = E[Y_n] \leq c \cdot n^3$$

$$\log 2^{E[X_n]} \leq \log(cn^3)$$

$$\underline{\underline{E[X_n]}} \leq \log c + \log n^3$$

$$\leq \underline{\underline{3 \log n + O(1)}}$$

we even get the leading constant!

It is known that $E[X_n] \approx 2.988$ [Devroye '86]

so this analysis was pretty good already.