Convexification by Averages

Iago Leal de Freitas
Advisor: Bernardo Freitas Paulo da Costa

Universidade Federal do Rio de Janeiro

November 11, 2019

Two-stage problems

Convex case

A two-stage **linear program** is

$$\begin{aligned} \min_x \quad & c_1^\top x + Q(x) \\ \text{s.t.} \quad & A_1 x = b_1 \\ & x \geqslant 0, \end{aligned}$$

where the cost-to-go function is

$$Q(x) = \begin{aligned} \min_y \quad & c^\top y \\ \text{s.t.} \quad & Ax + By = b \\ & y \geqslant 0. \end{aligned}$$

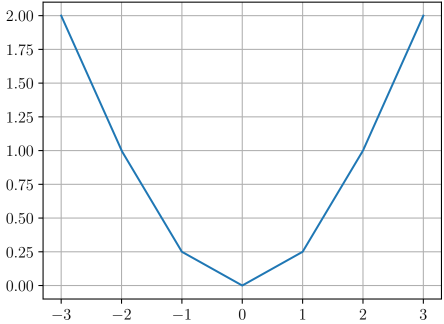


Figure: Example of cost-to-go.

Two-stage problems
Non-convex case

A two-stage **mixed integer linear program** is

$$\begin{array}{ll} \min\limits_{x} & c_1^\top x + Q(x) \\ \text{s.t.} & A_1 x = b_1 \\ & x \geqslant 0, \end{array}$$

where the cost-to-go function is

$$Q(x) = \begin{array}{ll} \min\limits_{y} & c^\top y \\ \text{s.t.} & Ax + By = b \\ & y \geqslant 0 \\ & y \in \mathbb{R}^n \times \mathbb{Z}^k. \end{array}$$

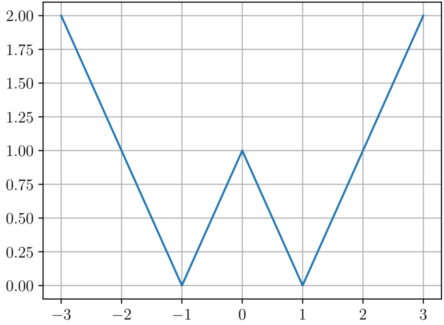


Figure: Example of cost-to-go.

Cost-to-go Q: **random function**.

- Uncertainty $\xi = (c^\top, A, B, b)$:

$$
\begin{aligned}
Q(x, \xi) = \quad \min_{y} \quad & c^\top y \\
\text{s.t.} \quad & Ax + By = b \\
& y \geqslant 0 \\
& y \in \mathbb{R}^n \times \mathbb{Z}^k.
\end{aligned}
$$

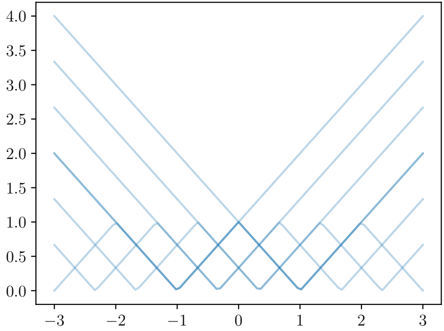


Figure: Random cost-to-go.

Cost-to-go Q: **random function**.

- First stage:

$$\begin{aligned} \min_x \quad & c_1^\top x + \mathbb{E}\left[Q(x,\xi)\right] \\ \text{s.t.} \quad & A_1 x = b_1 \\ & x \geqslant 0. \end{aligned}$$

- Uncertainty $\xi = (c^\top, A, B, b)$:

$$Q(x,\xi) = \begin{aligned} \min_y \quad & c^\top y \\ \text{s.t.} \quad & Ax + By = b \\ & y \geqslant 0 \\ & y \in \mathbb{R}^n \times \mathbb{Z}^k. \end{aligned}$$

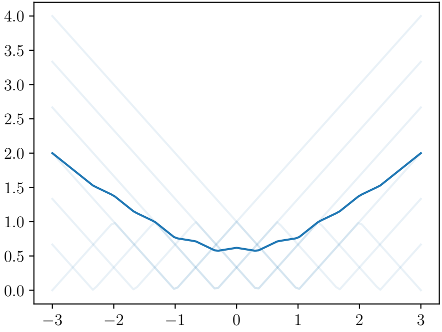


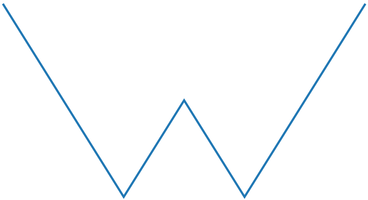
Figure: Expected cost-to-go.

Average of non-convex functions

Question

What is a characterization of the expected cost-to-go?

- Average of convex functions is **convex**.
- For **non-convex** functions:

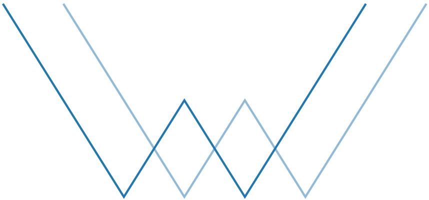


Average of non-convex functions

Question

What is a characterization of the expected cost-to-go?

- Average of convex functions is **convex**.
- For **non-convex** functions:

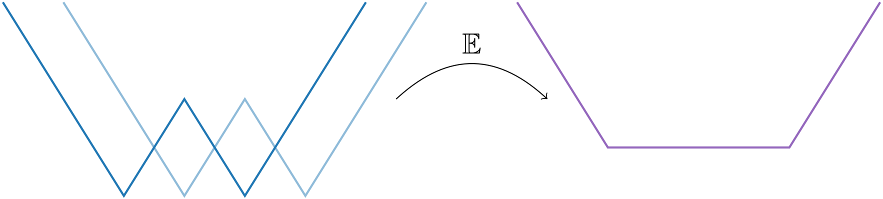


Average of non-convex functions

Question

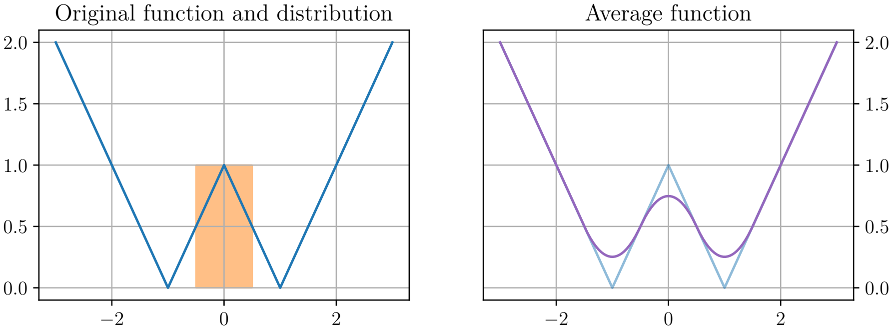
What is a characterization of the expected cost-to-go?

- Average of convex functions is **convex**.
- For **non-convex** functions: it may also be!



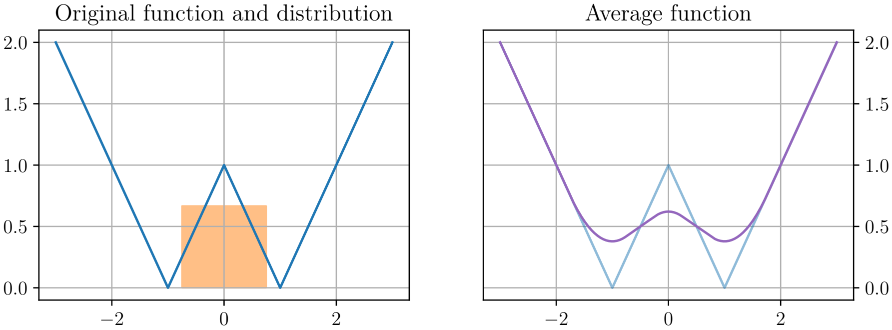
Non-convexity reduction

Uniform noises



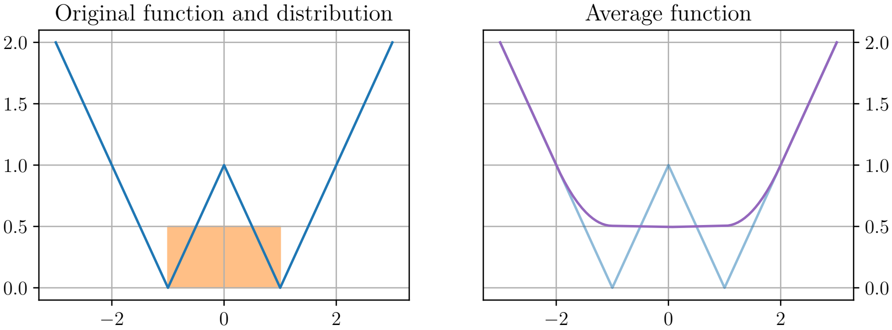
Non-convexity reduction

Uniform noises



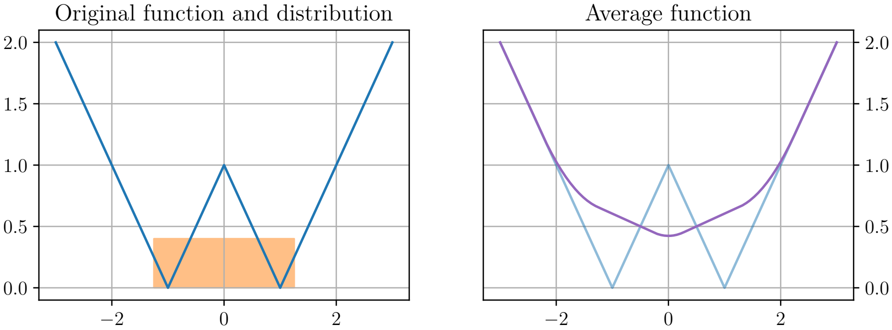
Non-convexity reduction

Uniform noises



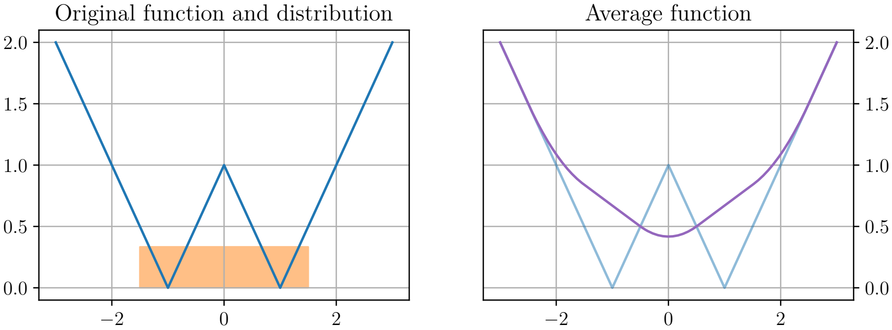
Non-convexity reduction

Uniform noises



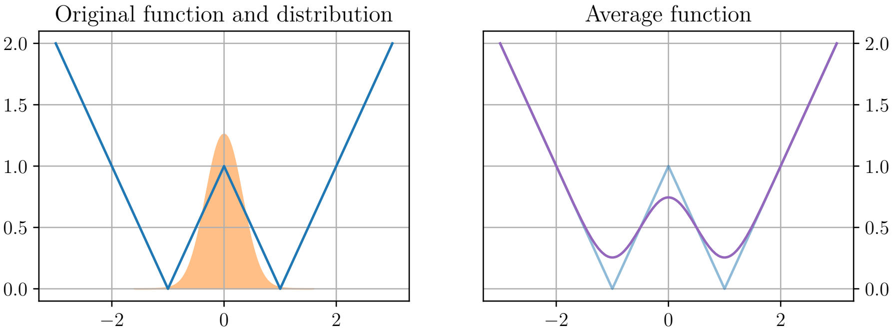
Non-convexity reduction

Uniform noises



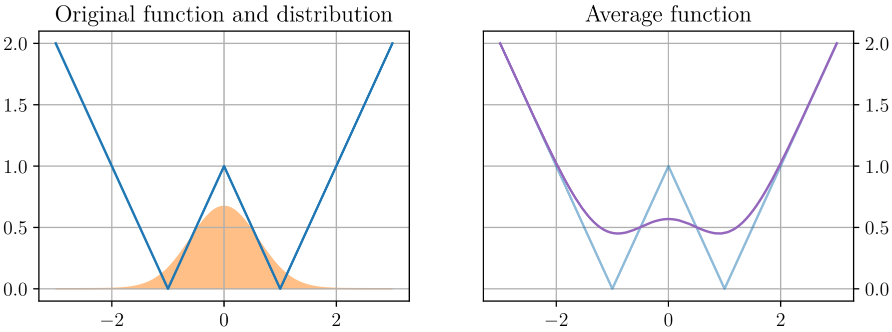
Non-convexity reduction

Gaussian noises



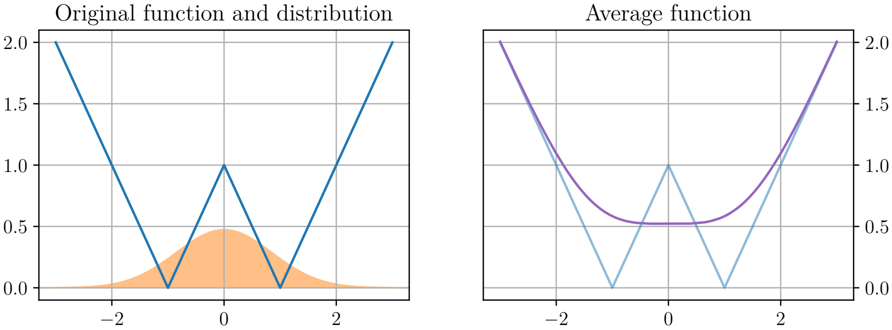
Non-convexity reduction

Gaussian noises



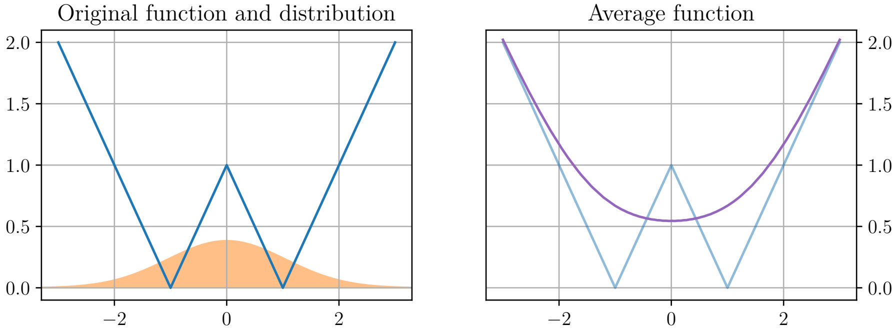
Non-convexity reduction

Gaussian noises



Non-convexity reduction

Gaussian noises



Measuring the non-convexity of a function

Question

How to rigorously measure the non-convexity of a function?

Measuring the non-convexity of a function

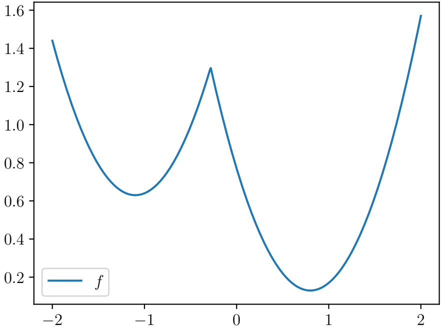
Question

How to rigorously measure the non-convexity of a function?

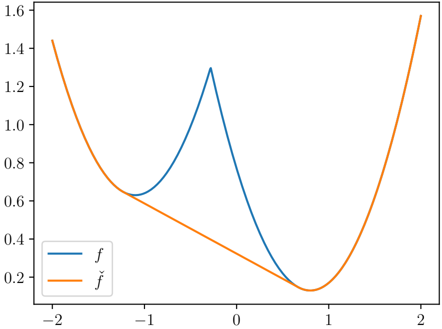
Idea

A natural approach is to consider the **gap** between it and its convex relaxation!

- Let f be a function.



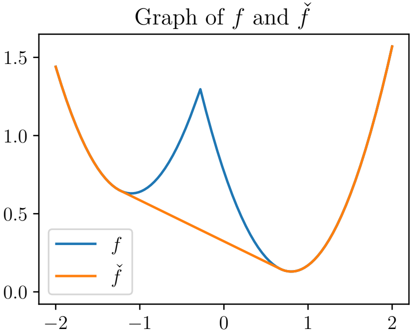
- Let f be a function.
- Its **convex relaxation** $\check{f}$ is the largest convex function everywhere less than f.

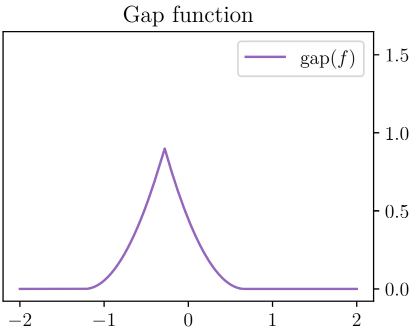


Measuring the non-convexity of a function

Gap function

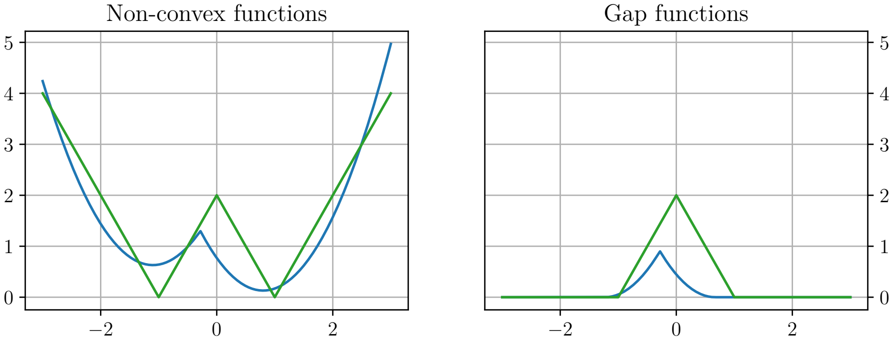
- Define: $\operatorname{gap}(f) = f - \breve{f}$;
- Notice: the gap is identically zero if and only if f is convex,
- $\operatorname{gap}(f)$ is always non-negative;





- We can say that f **is less non-convex than** g if

$$\operatorname{gap}(f) \leqslant \operatorname{gap}(g).$$



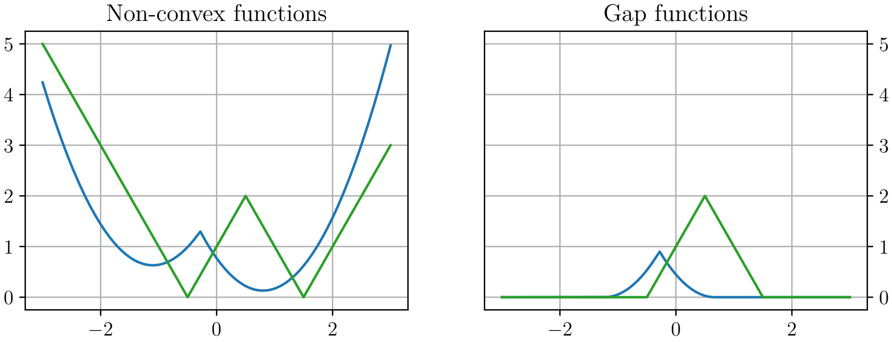
Measuring the non-convexity of a function

Gap comparisons

- We can say that f **is less non-convex than** g if

$$\operatorname{gap}(f) \leqslant \operatorname{gap}(g).$$

- But their gaps may not be comparable...



- Solution: project on $\mathbb{R}$ using a norm $\|\cdot\|$.
- Requirement: preserve gap comparisons,

$$\operatorname{gap}(f) \leqslant \operatorname{gap}(g) \implies \|\operatorname{gap}(f)\| \leqslant \|\operatorname{gap}(g)\|.$$

- Solution: project on $\mathbb{R}$ using a norm $\|\cdot\|$.
- Requirement: preserve gap comparisons,

$$\operatorname{gap}(f) \leqslant \operatorname{gap}(g) \implies \|\operatorname{gap}(f)\| \leqslant \|\operatorname{gap}(g)\| .$$

Definition

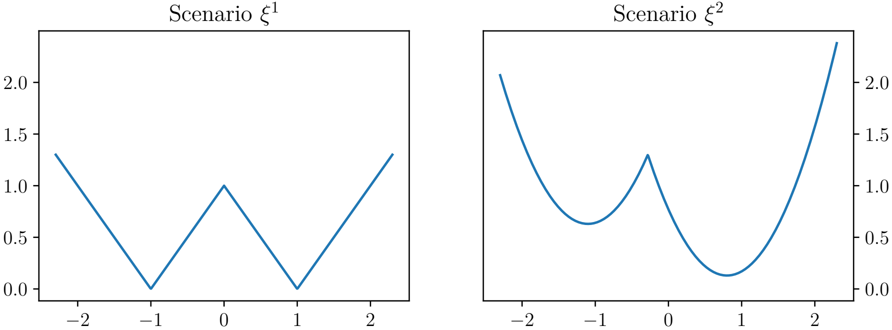
A function norm $\|\cdot\|$ is **monotone** if for all $g, h \geqslant 0$,

$$g \leqslant h \implies \|g\| \leqslant \|h\| .$$

Examples: The uniform norm, all p-norms...

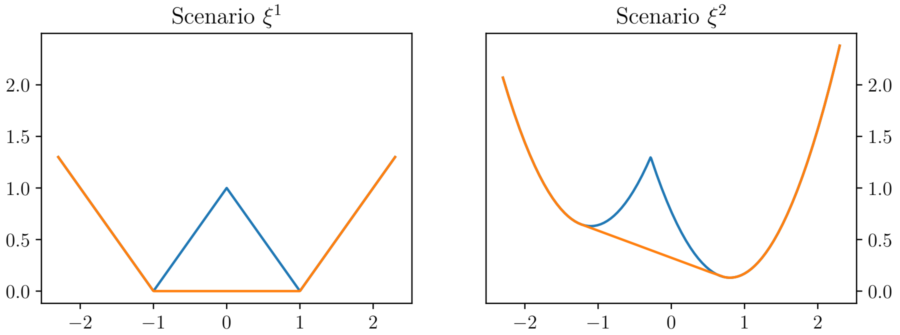
Interlude: notation for random functions

- Random variable: ξ.
- Random function: $Q(x, \xi)$.



Interlude: notation for random functions

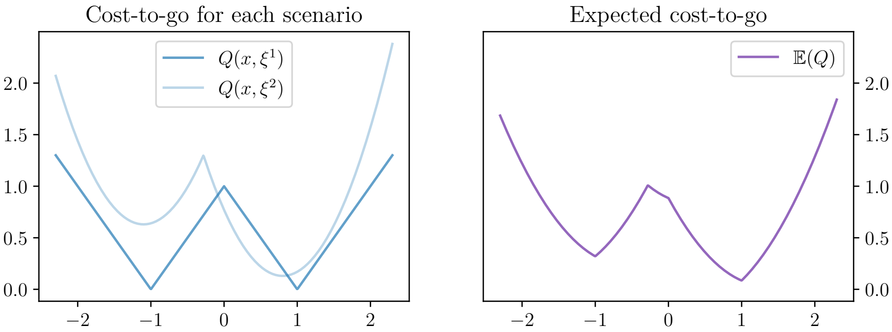
- Random variable: ξ.
- Random function: $Q(x, \xi)$.
- Convex relaxation $\check{Q}$: Defined for each realization $Q(\cdot, \xi)$.



Interlude: notation for random functions

- Random variable: ξ.
- Random function: $Q(x, \xi)$.
- Convex relaxation $\check{Q}$: Defined for each realization $Q(\cdot, \xi)$.
- Average function:

$$\mathbb{E}\left[Q\right] = x \mapsto \mathbb{E}^{\xi}\left[Q(x, \xi)\right]$$



Let Q be a random function (such as a cost-to-go function, for example)

- By definition: $\check{Q} \leqslant Q$.

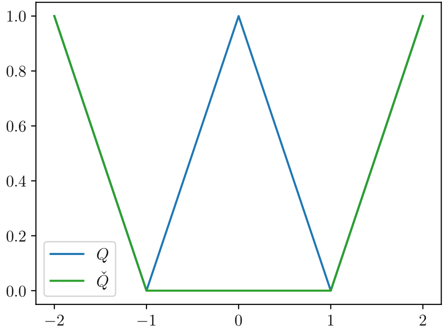


Figure: A realization of Q.

Let Q be a random function (such as a cost-to-go function, for example)

- By definition: $\check{Q} \leqslant Q$.
- Averages preserve inequalities:

$$\mathbb{E}\left[\check{Q}\right] \leqslant \mathbb{E}\left[Q\right].$$

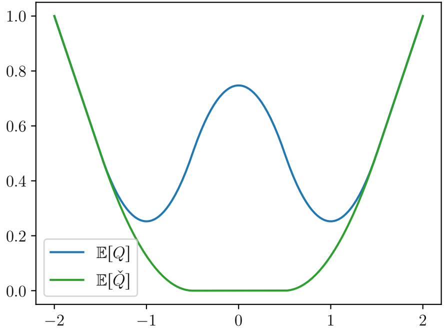


Figure: Average function.

Measuring the non-convexity of a function

Main inequality

Let Q be a random function (such as a cost-to-go function, for example)

- By definition: $\check{Q} \leqslant Q$.
- Averages preserve inequalities:

$$\mathbb{E}\left[\check{Q}\right] \leqslant \mathbb{E}\left[Q\right].$$

- By definition of convex relaxation:

$$\mathbb{E}\left[\check{Q}\right] \leqslant \widecheck{\mathbb{E}\left[Q\right]} \leqslant \mathbb{E}\left[Q\right].$$

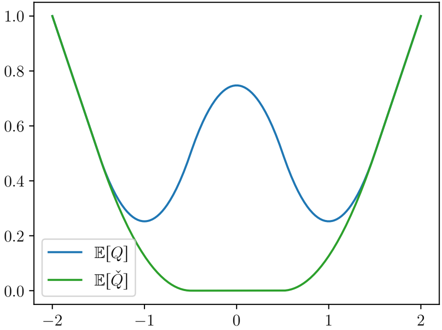


Figure: Average function.

Measuring the non-convexity of a function

Main inequality

Let Q be a random function (such as a cost-to-go function, for example)

- By definition: $\check{Q} \leqslant Q$.
- Averages preserve inequalities:

$$\mathbb{E}\left[\check{Q}\right] \leqslant \mathbb{E}\left[Q\right].$$

- By definition of convex relaxation:

$$\mathbb{E}\left[\check{Q}\right] \leqslant \widecheck{\mathbb{E}\left[Q\right]} \leqslant \mathbb{E}\left[Q\right].$$

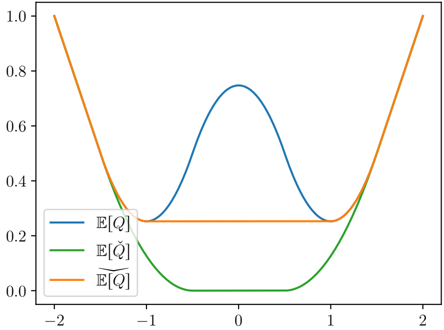


Figure: Average function.

$\mathbb{E}[\check{Q}]$ and $\widetilde{\mathbb{E}[Q]}$: underapproximations to $\mathbb{E}[Q]$.

- $\mathbb{E}[\check{Q}] \leqslant \widetilde{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q]$
- Rewriting them:

$$\mathbb{E}[Q] - \widetilde{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q] - \mathbb{E}[\check{Q}]$$

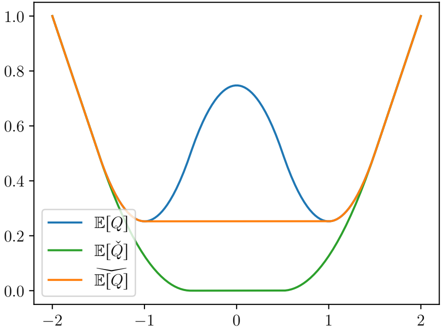


Figure: Average function.

$\mathbb{E}[\check{Q}]$ and $\widetilde{\mathbb{E}[Q]}$: underapproximations to $\mathbb{E}[Q]$.

- $\mathbb{E}[\check{Q}] \leqslant \widetilde{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q]$
- Rewriting them:

$$\begin{aligned} \mathbb{E}[Q] - \widetilde{\mathbb{E}[Q]} &\leqslant \mathbb{E}[Q] - \mathbb{E}[\check{Q}] \\ &= \mathbb{E}[Q - \check{Q}]. \end{aligned}$$

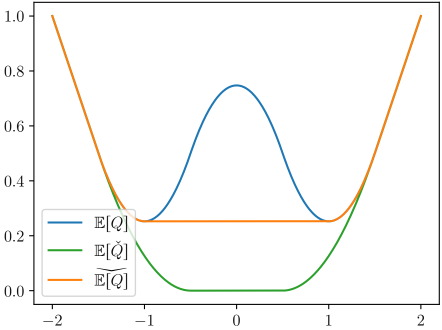


Figure: Average function.

Measuring the non-convexity of a function

Gap inequality

$\mathbb{E}[\check{Q}]$ and $\widetilde{\mathbb{E}[Q]}$: underapproximations to $\mathbb{E}[Q]$.

- $\mathbb{E}[\check{Q}] \leqslant \widetilde{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q]$
- Rewriting them:

$$\mathbb{E}[Q] - \widetilde{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q] - \mathbb{E}[\check{Q}]$$

- Expression for gap:

$$\operatorname{gap}(\mathbb{E}[Q]) \leqslant \mathbb{E}[\operatorname{gap}(Q)].$$

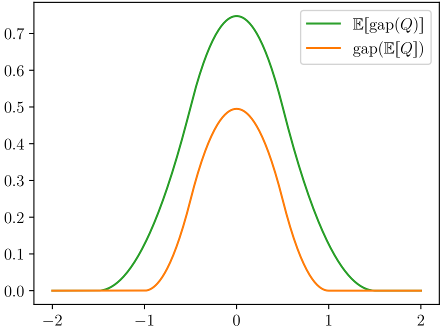


Figure: Gap functions.

Measuring the non-convexity of a function

Gap inequality

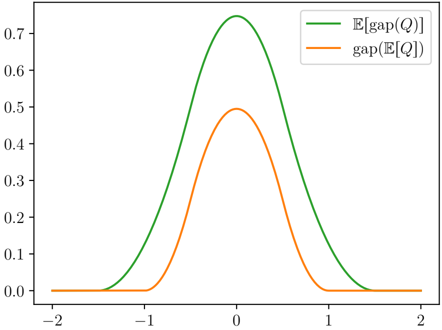
$\mathbb{E}[\check{Q}]$ and $\overbrace{\mathbb{E}[Q]}$: underapproximations to $\mathbb{E}[Q]$.

- $\mathbb{E}[\check{Q}] \leqslant \overbrace{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q]$
- Rewriting them:

$$\mathbb{E}[Q] - \overbrace{\mathbb{E}[Q]} \leqslant \mathbb{E}[Q] - \mathbb{E}[\check{Q}]$$

- Expression for gap:

$$\mathrm{gap}(\mathbb{E}[Q]) \leqslant \mathbb{E}[\mathrm{gap}(Q)].$$



- Monotone norm:

$$\|\mathrm{gap}(\mathbb{E}[Q])\| \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \leqslant \mathbb{E}\|\mathrm{gap}(Q)\|.$$

Back to optimization

- Common case: cost-to-go is only uncertain on stage transition,

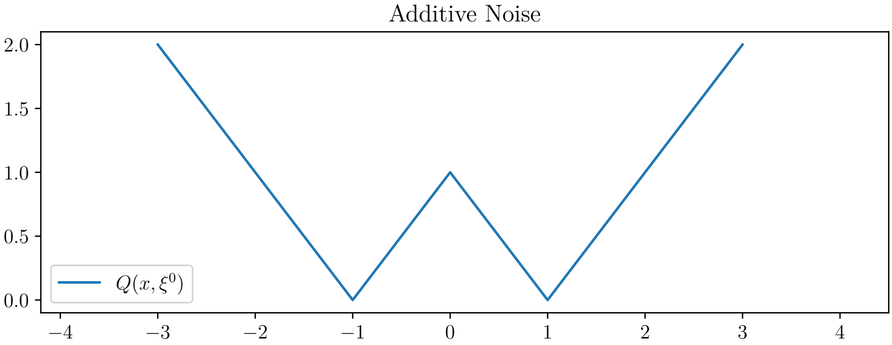
$$\begin{aligned} Q(x,\xi) = \min_y \quad & c^\top y \\ \text{s.t.} \quad & Ay = T(x - \xi) + b \\ & y \geqslant 0. \end{aligned}$$

Back to optimization

- Common case: cost-to-go is only uncertain on stage transition,

$$\begin{aligned} Q(x, \xi) = \min_y \quad & c^\top y \\ \text{s.t.} \quad & Ay = T(x - \xi) + b \\ & y \geqslant 0. \end{aligned}$$

- Additive noise: $Q(x, \xi) = f(x - \xi)$.

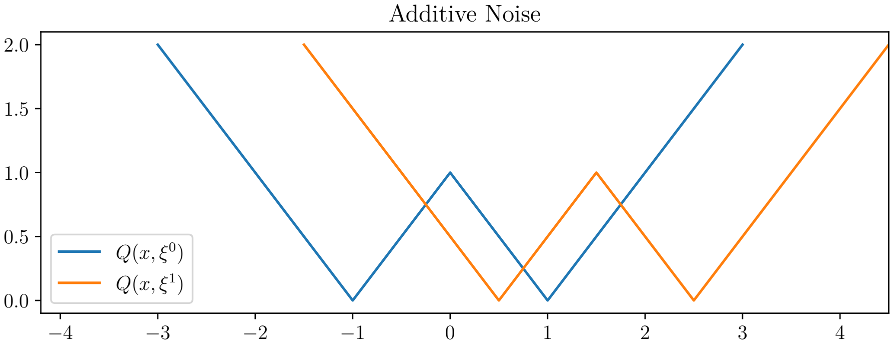


Back to optimization

- Common case: cost-to-go is only uncertain on stage transition,

$$\begin{aligned} Q(x,\xi) = \min_y \quad & c^\top y \\ \text{s.t.} \quad & Ay = T(x - \xi) + b \\ & y \geqslant 0. \end{aligned}$$

- Additive noise: $Q(x,\xi) = f(x - \xi)$.

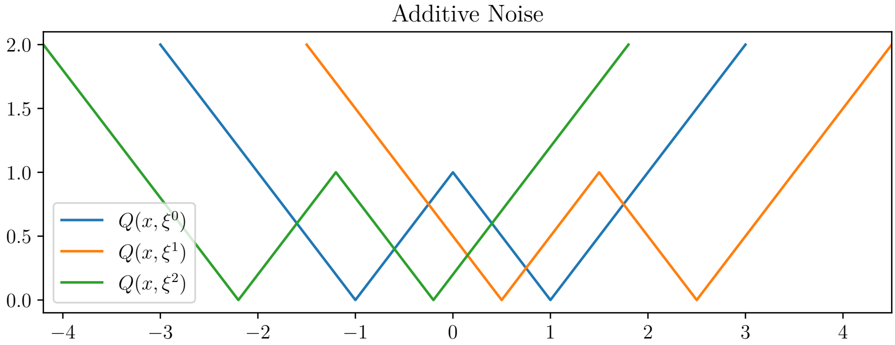


Back to optimization

- Common case: cost-to-go is only uncertain on stage transition,

$$\begin{aligned} Q(x,\xi) = \min_y \quad & c^\top y \\ \text{s.t.} \quad & Ay = T(x-\xi) + b \\ & y \geqslant 0. \end{aligned}$$

- Additive noise: $Q(x,\xi) = f(x-\xi)$.

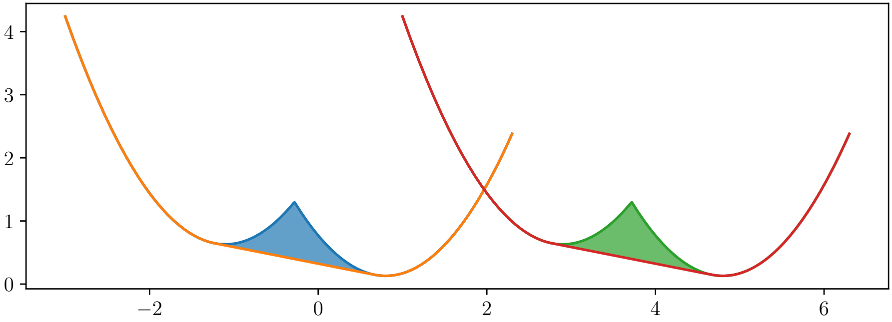


Additive noise

- Let $Q(x, \xi) = f(x - \xi)$.

Question

Is there a relation between the gap of $\mathbb{E}[Q]$ and the gap of f?



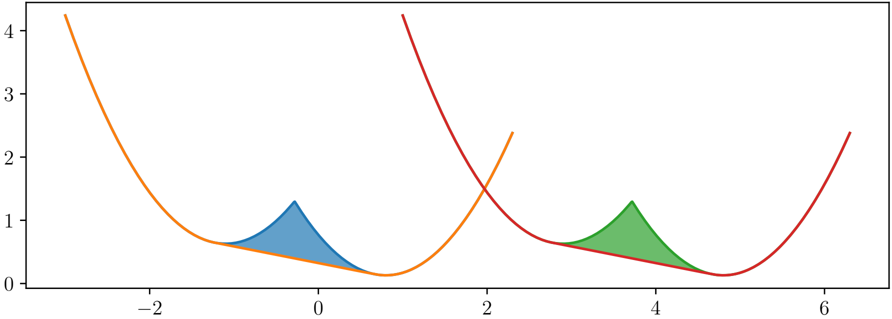
Additive noise

- Let $Q(x, \xi) = f(x - \xi)$.

Question

Is there a relation between the gap of $\mathbb{E}[Q]$ and the gap of f?

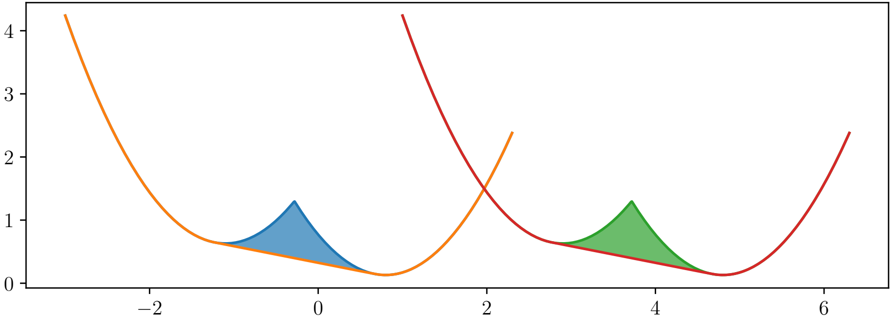
- Yes! With one condition: **translation invariance**.



Additive noise

- The **translation operator** τ_a shifts a function by a,

$$\tau_a f(x) = f(x - a).$$



Additive noise

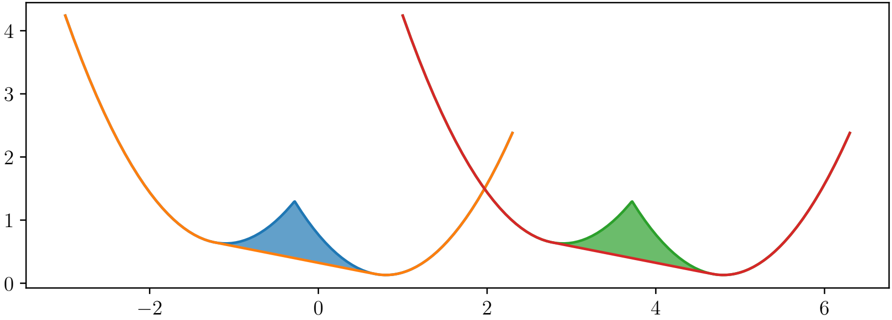
- The **translation operator** τ_a shifts a function by a,

$$\tau_a f(x) = f(x - a).$$

- A function norm is **translation invariant** if

$$\|f\| = \|\tau_a f\|.$$

Examples: Again, uniform norm and all p-norms.



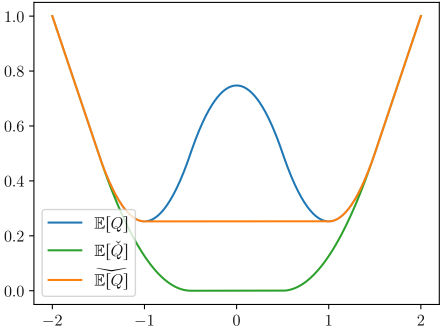
Additive noise

Inequality

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- From the previous inequalities:

$$\begin{aligned} \|\mathrm{gap}(\mathbb{E}[Q])\| &\leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \\ &\leqslant \mathbb{E}\|\mathrm{gap}(Q)\| \end{aligned}$$

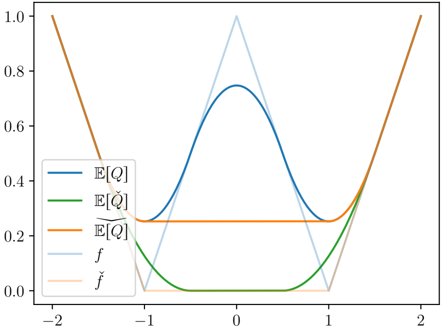


Additive noise
Inequality

- Let $Q(x,\xi) = \tau_\xi f(x) = f(x-\xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- From the previous inequalities:

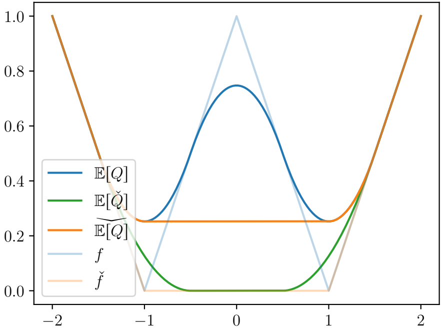
$$\begin{aligned}\|\mathrm{gap}(\mathbb{E}[Q])\| &\leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \\ &\leqslant \mathbb{E}\,\|\mathrm{gap}(Q)\| \\ &= \mathbb{E}\,\|\tau_\xi f - \widecheck{\tau_\xi f}\|\end{aligned}$$



- Let $Q(x,\xi) = \tau_\xi f(x) = f(x-\xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- From the previous inequalities:

$$
\begin{aligned}
\|\mathrm{gap}(\mathbb{E}[Q])\| &\leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \\
&\leqslant \mathbb{E}\,\|\mathrm{gap}(Q)\| \\
&= \mathbb{E}\,\|\tau_\xi f - \widecheck{\tau_\xi f}\| \\
&= \mathbb{E}\,\|\tau_\xi (f - \check{f})\|
\end{aligned}
$$



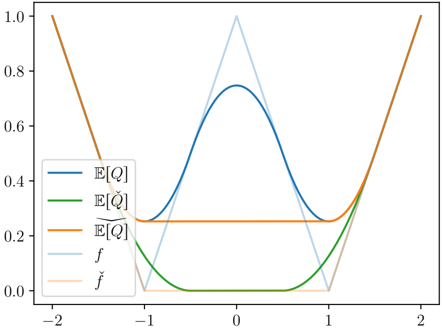
Additive noise

Inequality

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- From the previous inequalities:

$$
\begin{aligned}
\|\mathrm{gap}(\mathbb{E}[Q])\| &\leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \\
&\leqslant \mathbb{E}\,\|\mathrm{gap}(Q)\| \\
&= \mathbb{E}\,\|\tau_\xi f - \breve{\tau_\xi f}\| \\
&= \mathbb{E}\,\|\tau_\xi (f - \check{f})\| \\
&= \mathbb{E}\,\|f - \check{f}\|
\end{aligned}
$$

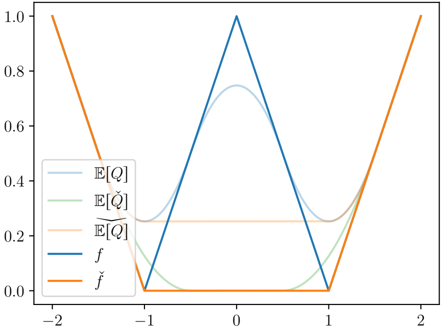


Additive noise
Inequality

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- From the previous inequalities:

$$
\begin{aligned}
\|\operatorname{gap}(\mathbb{E}[Q])\| &\leqslant \|\mathbb{E}[\operatorname{gap}(Q)]\| \\
&\leqslant \mathbb{E}\|\operatorname{gap}(Q)\| \\
&= \mathbb{E}\|\tau_\xi f - \widecheck{\tau_\xi f}\| \\
&= \mathbb{E}\|\tau_\xi (f - \check{f})\| \\
&= \mathbb{E}\|f - \check{f}\| \\
&= \|f - \check{f}\|
\end{aligned}
$$



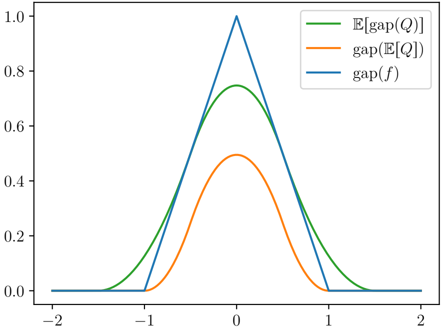
Additive noise

Inequality

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- From the previous inequalities:

$$
\begin{aligned}
\|\mathrm{gap}(\mathbb{E}[Q])\| &\leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \\
&\leqslant \mathbb{E}\|\mathrm{gap}(Q)\| \\
&= \mathbb{E}\|\tau_\xi f - \widecheck{\tau_\xi f}\| \\
&= \mathbb{E}\|\tau_\xi(f - \check{f})\| \\
&= \mathbb{E}\|f - \check{f}\| \\
&= \|f - \check{f}\| \\
&= \|\mathrm{gap}(f)\|
\end{aligned}
$$



- Inequality:

$$\|\mathrm{gap}(\mathbb{E}[Q])\| \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\| \leqslant \|\mathrm{gap}(f)\|.$$

- Inequality for uniform norm:

$$\|\mathrm{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\|_\infty \leqslant \|\mathrm{gap}(f)\|_\infty .$$

Additive noise

Uniform bounds

- Inequality for uniform norm:

$$\|\mathrm{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\|_\infty \leqslant \|\mathrm{gap}(f)\|_\infty .$$

- Improvement: probability of ξ being inside the support of $\mathrm{gap}(f)$ may be small,

Theorem

$$\|\mathrm{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\|_\infty \leqslant \kappa \|\mathrm{gap}(f)\|_\infty$$

where

$$\kappa = \sup_x \mathbb{P}\left[x - \xi \in \operatorname{supp} \mathrm{gap}(f)\right]$$

- Inequality for uniform norm:

$$\|\mathrm{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\|_\infty \leqslant \|\mathrm{gap}(f)\|_\infty .$$

- Improvement: probability of ξ being inside the support of $\mathrm{gap}(f)$ may be small,

Theorem

$$\|\mathrm{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\mathrm{gap}(Q)]\|_\infty \leqslant \kappa \|\mathrm{gap}(f)\|_\infty$$

where

$$\kappa = \sup_x \mathbb{P}\left[x - \xi \in \mathrm{supp}\,\mathrm{gap}(f)\right]$$

- κ is small if the distribution of ξ is “scattered enough”.

Theorem

If $\operatorname{gap}(f)$ *is* ***integrable*** *and* ξ_k *are random variables whose* ***densities*** μ_k *are bounded,*

$$\|\operatorname{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\operatorname{gap}(Q)]\|_\infty \leqslant \|\mu_k\|_\infty \|\operatorname{gap}(f)\|_1 .$$

Additive noise

Asymptotic behaviour

Theorem

If $\operatorname{gap}(f)$ *is* ***integrable*** *and* ξ_k *are random variables whose* ***densities*** μ_k *are bounded,*

$$\|\operatorname{gap}(\mathbb{E}[Q])\|_\infty \leqslant \|\mathbb{E}[\operatorname{gap}(Q)]\|_\infty \leqslant \|\mu_k\|_\infty \|\operatorname{gap}(f)\|_1 .$$

- This means that if

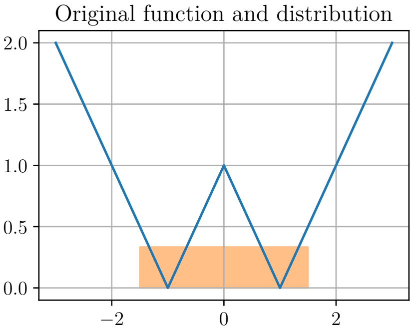
$$\|\mu_k\|_\infty \to 0,$$

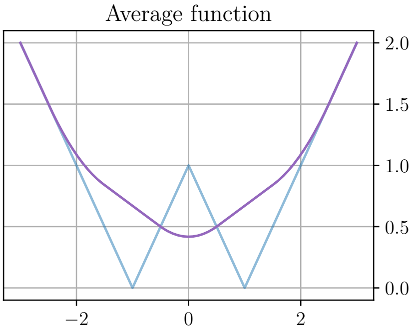
the expected function $\mathbb{E}[Q]$ becomes **asymptotically** convex.

Additive noise

Asymptotic behaviour

- $\xi_h \sim U[-h, h]$
- $\|\mu_h\|_\infty = \frac{1}{2h} \to 0$

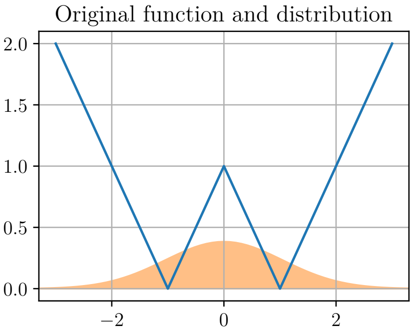


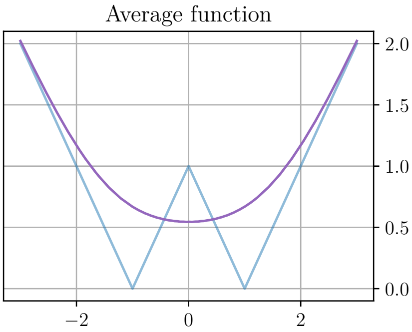


Additive noise

Asymptotic behaviour

- $\xi_h \sim N(m, \sigma^2)$
- $\|\mu_h\|_\infty = \frac{1}{\sqrt{2\pi\sigma^2}} \to 0$





Another approach: second derivative

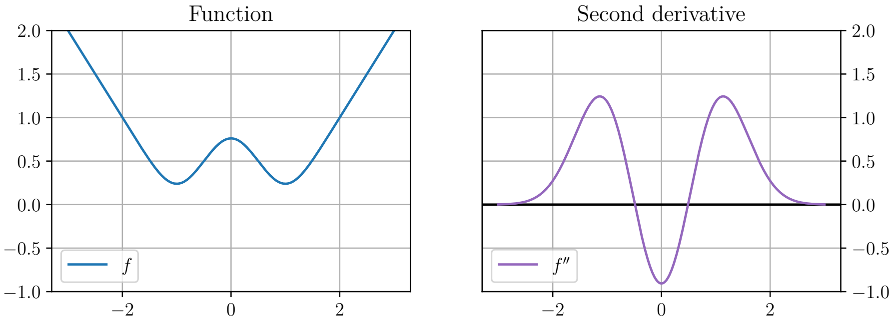
- Question: Why the gap?
- Famous result: $f\colon (a,b) \to \mathbb{R}$ twice differentiable,

$$f \text{ is convex} \iff f'' \geqslant 0.$$

Another approach: second derivative

- Question: Why the gap?
- Famous result: $f : (a, b) \to \mathbb{R}$ twice differentiable,

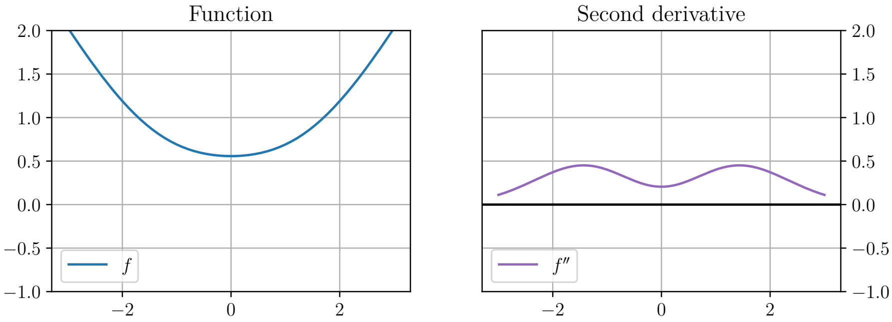
$$f \text{ is convex} \iff f'' \geqslant 0.$$



Another approach: second derivative

- Question: Why the gap?
- Famous result: $f : (a, b) \to \mathbb{R}$ twice differentiable,

$$f \text{ is convex} \iff f'' \geqslant 0.$$



Another approach: second derivative

Idea

Use the negative part of f'' to measure the non-convexity of f.

$$\mathcal{D}(f) := [f'']_-,$$

where $[x]_- = \max\{-x, 0\}$.

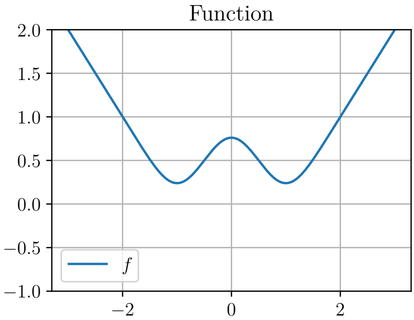
Another approach: second derivative

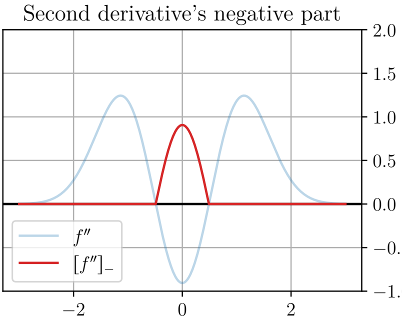
Idea

Use the negative part of f'' to measure the non-convexity of f.

$$\mathcal{D}(f) := [f'']_-,$$

where $[x]_- = \max\{-x, 0\}$.





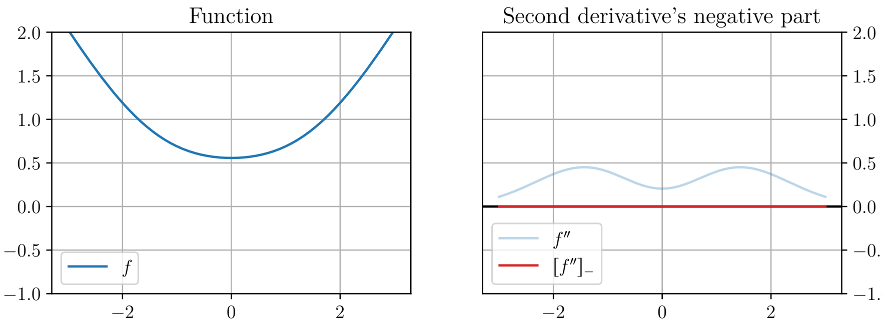
Another approach: second derivative

Idea

Use the negative part of f'' to measure the non-convexity of f.

$$\mathcal{D}(f) := [f'']_-,$$

where $[x]_- = \max\{-x, 0\}$.



Another approach: second derivative

Idea

Use the negative part of f'' to measure the non-convexity of f.

$$\mathcal{D}(f) := [f'']_-,$$

where $[x]_- = \max\{-x, 0\}$.

Properties:

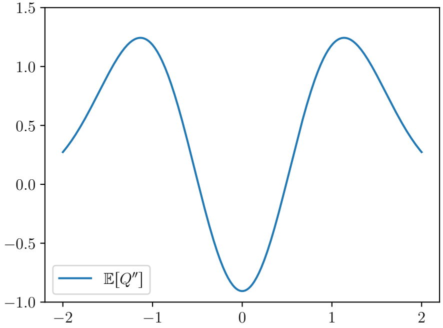
- $[f'']_-$ is identically zero if and only if f is convex,
- $[f'']_-$ is always non-negative.

Another approach: second derivative
Main inequality

Let Q be a smooth random function,

- Exchanging expected value and derivative:

$$\mathbb{E}\left[Q\right]'' = \mathbb{E}\left[Q''\right].$$



Another approach: second derivative

Main inequality

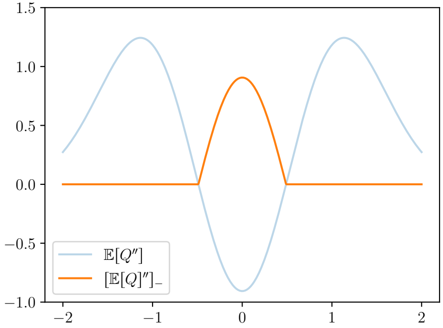
Let Q be a smooth random function,

- Exchanging expected value and derivative:

$$\mathbb{E}\left[Q\right]'' = \mathbb{E}\left[Q''\right].$$

- Convexity of negative part:

$$\max\left\{\mathbb{E}\left[Q''(x)\right], 0\right\} \leqslant \mathbb{E}\Big[\max\left\{Q''(x), 0\right\}\Big], \ \forall x.$$



Another approach: second derivative

Main inequality

Let Q be a smooth random function,

- Exchanging expected value and derivative:

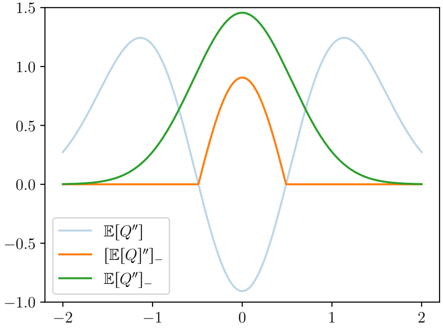
$$\mathbb{E}\left[Q\right]'' = \mathbb{E}\left[Q''\right].$$

- Convexity of negative part:

$$\max\left\{\mathbb{E}\left[Q''(x)\right], 0\right\} \leqslant \mathbb{E}\Big[\max\left\{Q''(x), 0\right\}\Big], \ \forall x.$$

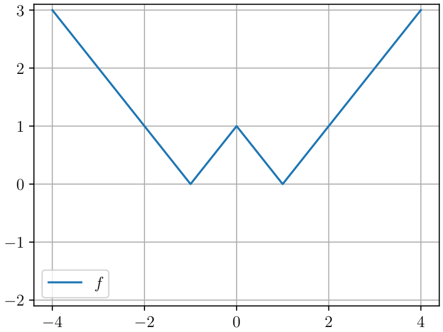
- Therefore:

$$[\mathbb{E}\left[Q\right]'']_- \leqslant \mathbb{E}\big[[Q'']_-\big].$$



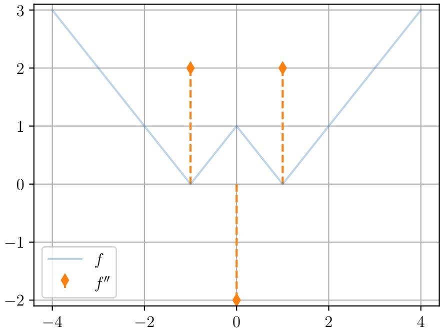
Derivative of cost-to-go

- The optimal value function of a mixed integer program is only **piecewise linear**.
- So its derivatives are not well-defined...



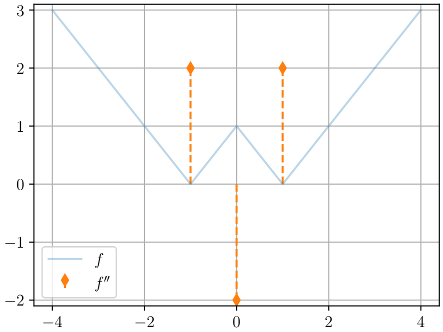
Derivative of cost-to-go

- The optimal value function of a mixed integer program is only **piecewise linear**.
- So its derivatives are not well-defined...
- Solution: **distributions**!



Derivative of cost-to-go

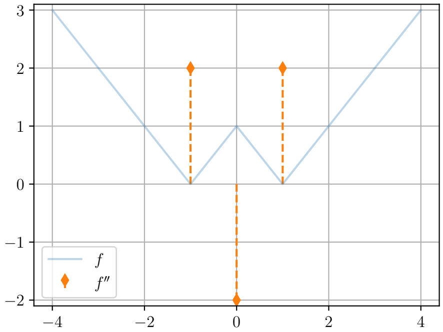
- The optimal value function of a mixed integer program is only **piecewise linear**.
- So its derivatives are not well-defined...
- Solution: **distributions**!



- Distribution theory: the second "derivative" of **any continuous function** is representable by a measure.

Derivative of cost-to-go

- The optimal value function of a mixed integer program is only **piecewise linear**.
- So its derivatives are not well-defined...
- Solution: **distributions**!



- Distribution theory: the second "derivative" of **any continuous function** is representable by a measure.
- More: f'' is a non-negative measure $\iff$ f is a convex function.

Hahn-Jordan decomposition

Every (signed) measure μ can be uniquely decomposed as the difference of two non-negative and mutually singular measures:

$$\mu = [\mu]_+ - [\mu]_-$$

Another approach: second derivative

Decomposition of measures

Hahn-Jordan decomposition

Every (signed) measure μ can be uniquely decomposed as the difference of two non-negative and mutually singular measures:

$$\mu = [\mu]_+ - [\mu]_-$$

Another non-convexity measure

Let $f\colon (a,b) \to \mathbb{R}$ be a continuous function. It is convex if and only if

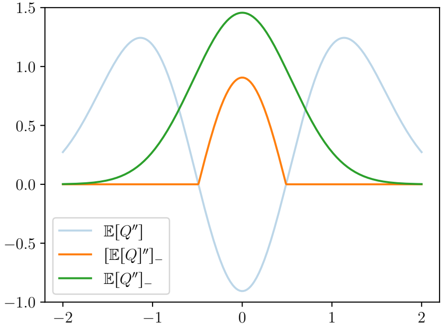
$$[f'']_- = 0.$$

Another approach: second derivative
Main inequality

Let Q be a random continuous function.

- It still holds that:

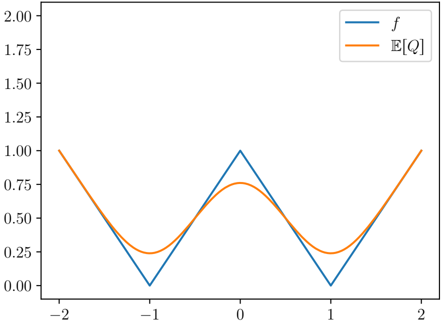
$$[\mathbb{E}[Q]'']_- \leqslant \mathbb{E}\big[[Q'']_-\big].$$



Reduction of non-convexity

Additive Noise

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

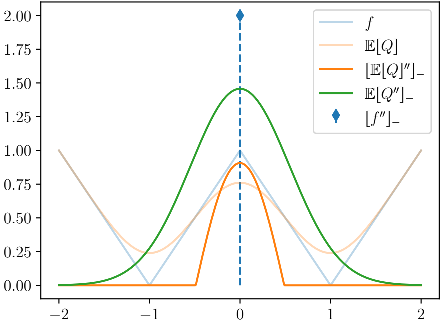


Reduction of non-convexity
Additive Noise

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

- f'' controls $\mathbb{E}\left[Q''\right]$:

$$\left\|\left[\mathbb{E}\left[Q\right]''\right]_-\right\| \leqslant \left\|\mathbb{E}\left[\left[Q''\right]_-\right]\right\| \leqslant \left\|\left[f''\right]_-\right\|.$$



Reduction of non-convexity

Additive Noise

- Let $Q(x, \xi) = \tau_\xi f(x) = f(x - \xi)$.
- $\|\cdot\|$: **monotone** and **translation invariant**.

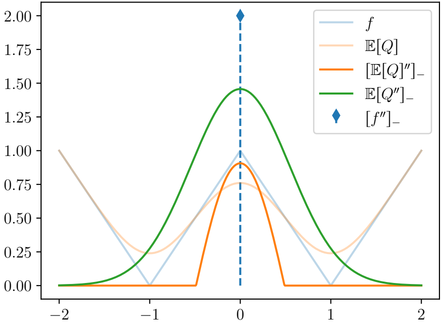
- f'' controls $\mathbb{E}\left[Q''\right]$:

$$\left\|\left[\mathbb{E}\left[Q\right]''\right]_-\right\| \leqslant \left\|\mathbb{E}\left[\left[Q''\right]_-\right]\right\| \leqslant \left\|\left[f''\right]_-\right\|.$$

- Uniform norm:

$$\left\|\mathbb{E}[Q'']_-\right\|_\infty \leqslant \kappa \left\|\left[f''\right]_-\right\|_\infty,$$

$$\kappa = \sup_x \mathbb{P}\left[x - \xi \in \operatorname{supp}[f'']_-\right].$$



Similarities between $\mathrm{gap}(f)$ and $[f'']_-$

Inequalities for $\mathrm{gap}(f)$:

- $\mathrm{gap}\left(\mathbb{E}\left[Q\right]\right) \leqslant \mathbb{E}\left[\mathrm{gap}(Q)\right]$,

Inequalities for $[f'']_-$:

- $\left[\mathbb{E}\left[Q\right]''\right]_- \leqslant \mathbb{E}\big[[Q'']_-\big]$

Similarities between $\mathrm{gap}(f)$ and $[f'']_-$

Inequalities for $\mathrm{gap}(f)$:

- $\mathrm{gap}\left(\mathbb{E}\left[Q\right]\right) \leqslant \mathbb{E}\left[\mathrm{gap}(Q)\right]$,
- Additive noise:

$$\left\|\mathbb{E}\left[\mathrm{gap}(Q)\right]\right\| \leqslant \left\|\mathrm{gap}(f)\right\|,$$

Inequalities for $[f'']_-$:

- $\left[\mathbb{E}\left[Q\right]''\right]_- \leqslant \mathbb{E}\big[[Q'']_-\big]$
- Additive noise:

$$\Big\|\mathbb{E}[Q'']_-\Big\| \leqslant \Big\|[f'']_-\Big\|,$$

Similarities between $\mathrm{gap}(f)$ and $[f'']_-$

Inequalities for $\mathrm{gap}(f)$:

- $\mathrm{gap}\left(\mathbb{E}\left[Q\right]\right) \leqslant \mathbb{E}\left[\mathrm{gap}(Q)\right]$,
- Additive noise:

$$\left\|\mathbb{E}\left[\mathrm{gap}(Q)\right]\right\| \leqslant \left\|\mathrm{gap}(f)\right\|,$$

- Uniform norm:

$$\left\|\mathbb{E}\left[\mathrm{gap}(Q)\right]\right\|_\infty \leqslant \kappa \left\|\mathrm{gap}(f)\right\|_\infty,$$

Inequalities for $[f'']_-$:

- $\left[\mathbb{E}\left[Q\right]''\right]_- \leqslant \mathbb{E}\big[[Q'']_-\big]$
- Additive noise:

$$\Big\|\mathbb{E}[Q'']_-\Big\| \leqslant \Big\|[f'']_-\Big\|,$$

- Uniform norm:

$$\Big\|\mathbb{E}[Q'']_-\Big\|_\infty \leqslant \kappa \Big\|[f'']_-\Big\|_\infty,$$

Similarities between $\text{gap}(f)$ and $[f'']_-$

Inequalities for $\text{gap}(f)$:

- $\text{gap}\left(\mathbb{E}\left[Q\right]\right) \leqslant \mathbb{E}\left[\text{gap}(Q)\right]$,
- Additive noise:

$$\left\|\mathbb{E}\left[\text{gap}(Q)\right]\right\| \leqslant \left\|\text{gap}(f)\right\|,$$

- Uniform norm:

$$\left\|\mathbb{E}\left[\text{gap}(Q)\right]\right\|_\infty \leqslant \kappa \left\|\text{gap}(f)\right\|_\infty,$$

- Asymptotic:

$$\left\|\mathbb{E}\left[\text{gap}(Q)\right]\right\|_\infty \xrightarrow{\|\mu_k\|_\infty \to 0} 0.$$

Inequalities for $[f'']_-$:

- $\left[\mathbb{E}\left[Q\right]''\right]_- \leqslant \mathbb{E}\left[[Q'']_-\right]$
- Additive noise:

$$\left\|\mathbb{E}[Q'']_-\right\| \leqslant \left\|[f'']_-\right\|,$$

- Uniform norm:

$$\left\|\mathbb{E}[Q'']_-\right\|_\infty \leqslant \kappa \left\|[f'']_-\right\|_\infty,$$

- Asymptotic:

$$\left\|\mathbb{E}[Q'']_-\right\|_\infty \xrightarrow{\|\mu_k\|_\infty \to 0} 0.$$

Convex cones

Definition

A **pointed convex cone** K is a set satisfying

- Containing all rays:

$$\forall \lambda \geqslant 0,\, x \in K \implies \lambda x \in K.$$

- Convexity:

$$x,\, y \in K \implies x + y \in K,$$

- Containing no lines:

$$x \in K \text{ and } -x \in K \implies x = 0.$$

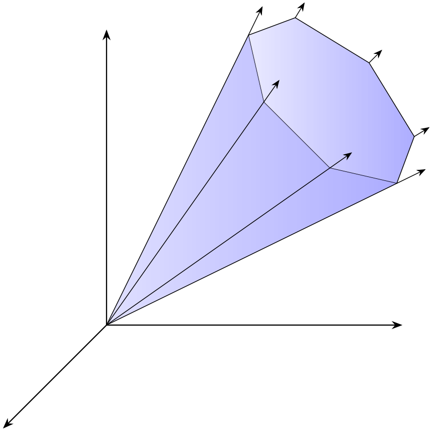


Figure: Example of convex cone.

Convexity with respect to a cone

Definition

Every proper convex cone K induces a **partial order** $\leqslant_K$ compatible with the linear structure and given by

$$a \leqslant_K b \iff b - a \in K.$$

Convexity with respect to a cone

Definition

Every proper convex cone K induces a **partial order** $\leqslant_K$ compatible with the linear structure and given by

$$a \leqslant_K b \iff b - a \in K.$$

Definition

A function $f\colon X \to Y$ is **convex with respect to a cone** $K \subset Y$ if its domain is a convex set and for all $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \leqslant_K \lambda f(x) + (1 - \lambda) f(y).$$

Non-convexity measures

Definition

A **non-convexity measure** on a set X of functions is an operator $\mathcal{M}\colon X \to Y$ whose codomain has a conic order $\leqslant_K$ such that

- $\mathcal{M}(f) = 0 \iff f$ is convex;
- $\mathcal{M}(f) \geqslant_K 0, \forall f$;
- $\mathcal{M}$ is convex with respect to K.

$\mathcal{M} = \text{gap}$ with K the cone of non-negative functions:

$\mathcal{M} = f \mapsto [f'']_-$ with K the cone of non-negative measures:

Non-convexity measures

Definition

A **non-convexity measure** on a set X of functions is an operator $\mathcal{M}\colon X \to Y$ whose codomain has a conic order $\leqslant_K$ such that

- $\mathcal{M}(f) = 0 \iff f$ is convex;
- $\mathcal{M}(f) \geqslant_K 0, \; \forall f$;
- $\mathcal{M}$ is convex with respect to K.

$\mathcal{M} = \mathrm{gap}$ with K the cone of non-negative functions:

- $\mathrm{gap}(f) = 0 \iff f$ is convex,

$\mathcal{M} = f \mapsto [f'']_-$ with K the cone of non-negative measures:

- $[f'']_- = 0 \iff f$ is convex,

Non-convexity measures

Definition

A **non-convexity measure** on a set X of functions is an operator $\mathcal{M}\colon X \to Y$ whose codomain has a conic order $\leqslant_K$ such that

- $\mathcal{M}(f) = 0 \iff f$ is convex;
- $\mathcal{M}(f) \geqslant_K 0$, $\forall f$;
- $\mathcal{M}$ is convex with respect to K.

$\mathcal{M} = \mathrm{gap}$ with K the cone of non-negative functions:

- $\mathrm{gap}(f) = 0 \iff f$ is convex,
- $\mathrm{gap}(f)$ is non-negative function,

$\mathcal{M} = f \mapsto [f'']_-$ with K the cone of non-negative measures:

- $[f'']_- = 0 \iff f$ is convex,
- $[f'']_-$ is non-negative measure,

Non-convexity measures

Definition

A **non-convexity measure** on a set X of functions is an operator $\mathcal{M}\colon X \to Y$ whose codomain has a conic order $\leqslant_K$ such that

- $\mathcal{M}(f) = 0 \iff f$ is convex;
- $\mathcal{M}(f) \geqslant_K 0$, $\forall f$;
- $\mathcal{M}$ is convex with respect to K.

$\mathcal{M} = \mathrm{gap}$ with K the cone of non-negative functions:

- $\mathrm{gap}(f) = 0 \iff f$ is convex,
- $\mathrm{gap}(f)$ is non-negative function,
- $\mathrm{gap}(\mathbb{E}\,[Q]) \leqslant \mathbb{E}\,[\mathrm{gap}(Q)]$.

$\mathcal{M} = f \mapsto [f'']_-$ with K the cone of non-negative measures:

- $[f'']_- = 0 \iff f$ is convex,
- $[f'']_-$ is non-negative measure,
- $[(\mathbb{E}Q)'']_- \leqslant \mathbb{E}[Q'']_-$.

Properties of non-convexity measures

- Jensen's inequality:

$$\mathcal{M}(\mathbb{E}\left[Q\right]) \leqslant \mathbb{E}\left[\mathcal{M}(Q)\right],$$

Properties of non-convexity measures

- Jensen's inequality:

$$\mathcal{M}(\mathbb{E}[Q]) \leqslant \mathbb{E}[\mathcal{M}(Q)],$$

- Preserved by monotone norms:
  If $\|\cdot\|$ is K-monotone, $\|\cdot\| \circ \mathcal{M}$ is also a non-convexity measure.

Properties of non-convexity measures

- Jensen's inequality:

$$\mathcal{M}(\mathbb{E}[Q]) \leqslant \mathbb{E}[\mathcal{M}(Q)],$$

- Preserved by monotone norms:
  If $\|\cdot\|$ is K-monotone, $\|\cdot\| \circ \mathcal{M}$ is also a non-convexity measure.
- Translation invariance: $(\, Q(x,\xi) = \tau_\xi f \,)$

$$\forall a,\; \mathcal{M}(\tau_a f) = \mathcal{M}(f) \implies \mathbb{E}[\mathcal{M}(Q)] \leqslant \mathcal{M}(f).$$

Properties of non-convexity measures

- Jensen's inequality:

$$\mathcal{M}(\mathbb{E}[Q]) \leqslant \mathbb{E}[\mathcal{M}(Q)],$$

- Preserved by monotone norms:
  If $\|\cdot\|$ is K-monotone, $\|\cdot\| \circ \mathcal{M}$ is also a non-convexity measure.
- Translation invariance: $(Q(x,\xi) = \tau_\xi f)$

$$\forall a,\, \mathcal{M}(\tau_a f) = \mathcal{M}(f) \implies \mathbb{E}[\mathcal{M}(Q)] \leqslant \mathcal{M}(f).$$

- Uniform bounds:

If $\mathcal{M}(f)$ is bounded,

$$\|\mathbb{E}[\mathcal{M}(Q)]\|_\infty \leqslant \kappa \|\mathcal{M}(f)\|_\infty$$

where
$\kappa = \sup_x \mathbb{P}\left[x - \xi \in \operatorname{supp} \mathcal{M}(f)\right]$.

If $\mathcal{M}(f)$ is integrable,

$$\|\mathbb{E}[\mathcal{M}(Q)]\|_\infty \leqslant \|\mu\|_\infty \|\mathcal{M}(f)\|_1$$

where μ is the probability density of the random variable.

- Second derivative for $\mathbb{R}^n$,
- f is convex $\iff$ its Hessian $D^2 f$ is positive semi-definite.

Non-convexity measure

Negative part of smallest eigenvalue of $D^2 f$,

$$\mathcal{R}(f) = [\lambda_1(D^2 f)]_-.$$

- Second derivative for $\mathbb{R}^n$,
- f is convex $\iff$ its Hessian $D^2 f$ is positive semi-definite.

Non-convexity measure

Negative part of smallest eigenvalue of $D^2 f$,

$$\mathcal{R}(f) = [\lambda_1(D^2 f)]_-.$$

- $\mathcal{R}(f) = 0 \iff f$ is convex.

- Second derivative for $\mathbb{R}^n$,
- f is convex $\iff$ its Hessian $D^2 f$ is positive semi-definite.

Non-convexity measure

Negative part of smallest eigenvalue of $D^2 f$,

$$\mathcal{R}(f) = [\lambda_1(D^2 f)]_-.$$

- $\mathcal{R}(f) = 0 \iff f$ is convex.
- $\mathcal{R}(f) \geqslant 0$.

Other examples

Hessian's smallest eigenvalue

- Second derivative for $\mathbb{R}^n$,
- f is convex $\iff$ its Hessian $D^2 f$ is positive semi-definite.

Non-convexity measure

Negative part of smallest eigenvalue of $D^2 f$,

$$\mathcal{R}(f) = [\lambda_1(D^2 f)]_-.$$

- $\mathcal{R}(f) = 0 \iff f$ is convex.
- $\mathcal{R}(f) \geqslant 0$.
- Min-max theorem:

$$\lambda_1(D^2 f) = \inf_{\|v\|_2 = 1} v^t (D^2 f) v.$$

Risk-averse convexification

Coherent risk measures

- Expected value is too "neutral".
- Sometimes, we want to be risk-averse.
- Solution: Exchange $\mathbb{E}$ in all formulas for a **risk measure**.

- Expected value is too "neutral".
- Sometimes, we want to be risk-averse.
- Solution: Exchange $\mathbb{E}$ in all formulas for a **risk measure**.

Definition

A **coherent risk measure** is a function ρ satisfying:

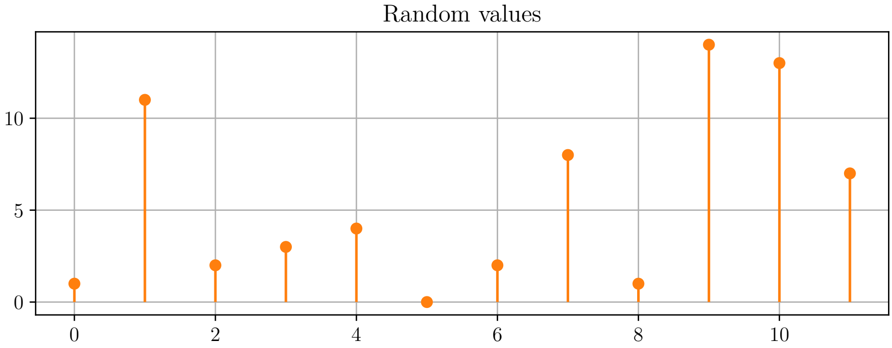
- Monotonicity: $X \leqslant Y \implies \rho(X) \leqslant \rho(Y)$;
- Translation equivariance: for all $a \in \mathbb{R}$, $\rho(X + a) = \rho(X) + a$;
- Convexity: for all $\lambda \in [0, 1]$,

$$\rho(\lambda X + (1 - \lambda)Y) \leqslant \lambda\rho(X) + (1 - \lambda)\rho(Y);$$

- Positive homogeneity: for all $t \geqslant 0$, $\rho(tX) = t\rho(X)$.

Coherent risk measures

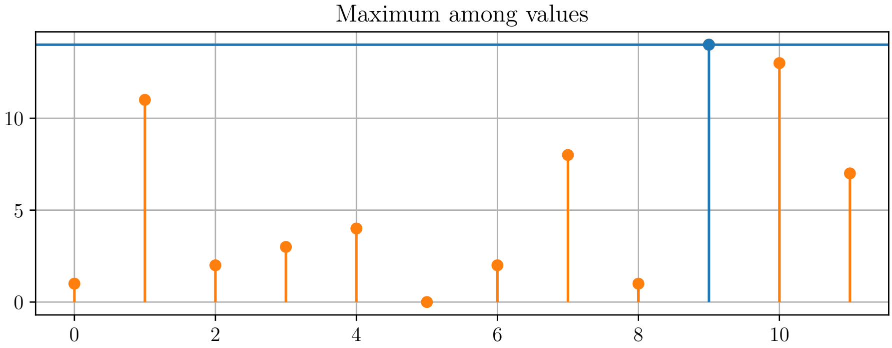
Examples of coherent risk measures:



Coherent risk measures

Examples of coherent risk measures:

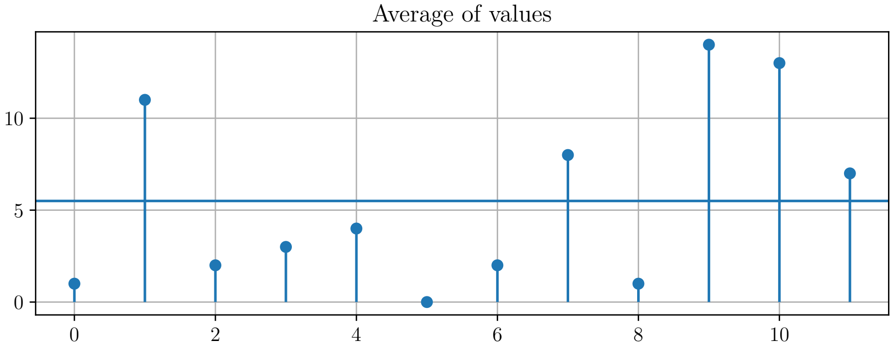
- Maximum value: $\sup X$;



Coherent risk measures

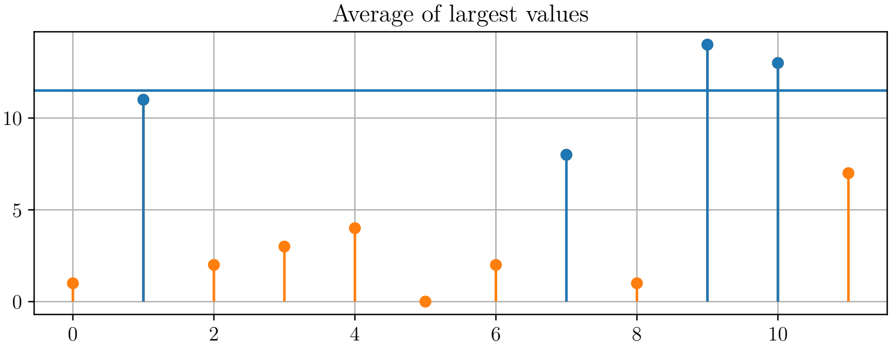
Examples of coherent risk measures:

- Maximum value: $\sup X$;
- Expected value: $\mathbb{E}[X]$;



Examples of coherent risk measures:

- Maximum value: $\sup X$;
- Expected value: $\mathbb{E}[X]$;
- Expectation of largest values;



Coherent risk measures

Examples of coherent risk measures:

- Maximum value: $\sup X$;
- Expected value: $\mathbb{E}[X]$;
- Expectation of largest values;

Dual representation

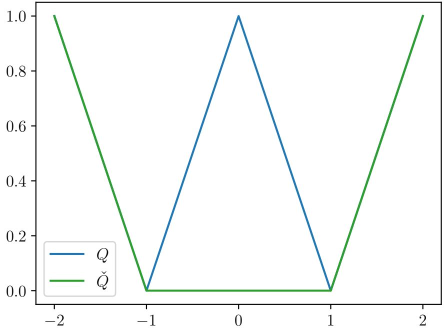
A coherent risk measure can be written as

$$\rho(X) = \sup_{\mu \in \mathcal{P}} \mathbb{E}^{\mu}[X]$$

for a given family of probabilities $\mathcal{P}$.

Let Q be a random function, ρ a coherent risk measure.

- By definition: $\check{Q} \leqslant Q$.



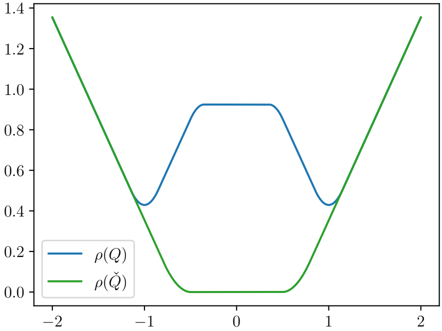
Risk-averse convexification

Main inequality

Let Q be a random function, ρ a coherent risk measure.

- By definition: $\check{Q} \leqslant Q$.
- Monotonicity:

$$\rho(\check{Q}) \leqslant \rho(Q).$$



Risk-averse convexification

Main inequality

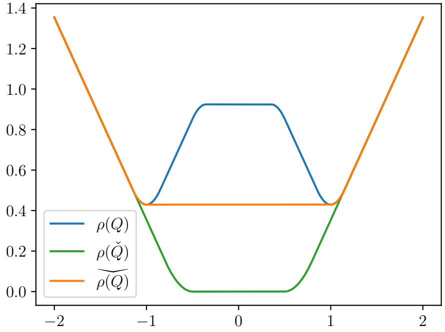
Let Q be a random function, ρ a coherent risk measure.

- By definition: $\check{Q} \leqslant Q$.
- Monotonicity:

$$\rho(\check{Q}) \leqslant \rho(Q).$$

- Convex relaxation:

$$\rho(\check{Q}) \leqslant \widetilde{\rho(Q)} \leqslant \rho(Q).$$

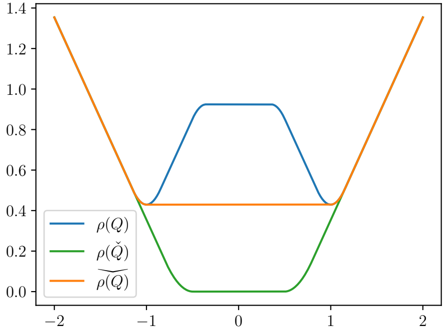


Risk-averse convexification

Gap inequality

$\rho(\check{Q})$ and $\widetilde{\rho(Q)}$: underapproximations to $\rho(Q)$.

- $\rho(\check{Q}) \leqslant \widetilde{\rho(Q)} \leqslant \rho(Q)$



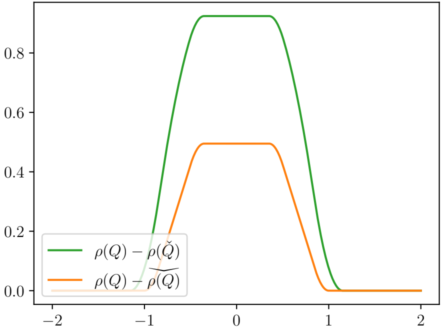
Risk-averse convexification

Gap inequality

$\rho(\check{Q})$ and $\overbrace{\rho(Q)}$: underapproximations to $\rho(Q)$.

- $\rho(\check{Q}) \leqslant \overbrace{\rho(Q)} \leqslant \rho(Q)$
- Rewriting them:

$$\rho(Q) - \overbrace{\rho(Q)} \leqslant \rho(Q) - \rho(\check{Q}).$$



Risk-averse convexification

Gap inequality

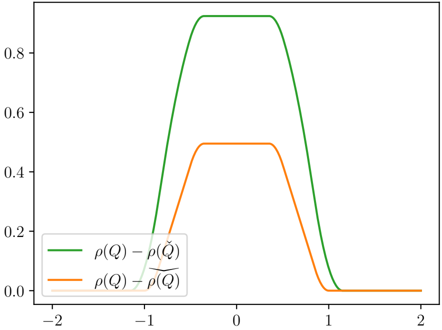
$\rho(\check{Q})$ and $\widetilde{\rho(Q)}$: underapproximations to $\rho(Q)$.

- $\rho(\check{Q}) \leqslant \widetilde{\rho(Q)} \leqslant \rho(Q)$
- Rewriting them:

$$\rho(Q) - \widetilde{\rho(Q)} \leqslant \rho(Q) - \rho(\check{Q}).$$

- Subadditivity:

$$\begin{aligned}\rho(Q) &= \rho(Q + \check{Q} - \check{Q}) \\ &\leqslant \rho(Q - \check{Q}) + \rho(\check{Q}).\end{aligned}$$



Risk-averse convexification

Gap inequality

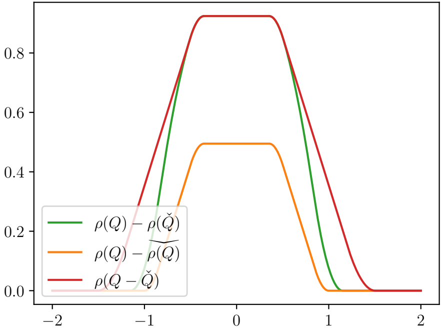
$\rho(\check{Q})$ and $\widetilde{\rho(Q)}$: underapproximations to $\rho(Q)$.

- $\rho(\check{Q}) \leqslant \widetilde{\rho(Q)} \leqslant \rho(Q)$
- Rewriting them:

$$\rho(Q) - \widetilde{\rho(Q)} \leqslant \rho(Q) - \rho(\check{Q}).$$

- Subadditivity:

$$\begin{aligned} \rho(Q) &= \rho(Q + \check{Q} - \check{Q}) \\ &\leqslant \rho(Q - \check{Q}) + \rho(\check{Q}). \end{aligned}$$



- Putting it all together:

$$\rho(Q) - \widetilde{\rho(Q)} \leqslant \rho(Q) - \rho(\check{Q}) \leqslant \rho(Q - \check{Q}).$$

Question

When the uncertainty is additive,

$$Q(x, \xi) = f(x - \xi) = \tau_\xi f(x),$$

the results regarding translation invariant norms still hold?

Question

When the uncertainty is additive,

$$Q(x, \xi) = f(x - \xi) = \tau_\xi f(x),$$

the results regarding translation invariant norms still hold?

Answer

Not in general...

Additive noise

Counterexample

- Expected value: (always)

$$\|\mathrm{gap}(\mathbb{E}\left[Q\right])\| \leqslant \|\mathbb{E}\left[\mathrm{gap}(Q)\right]\| \leqslant \|\mathrm{gap}(f)\|\,.$$

- Risk measures: (always)

$$\|\mathrm{gap}(\rho(Q))\| \leqslant \|\rho(\mathrm{gap}(Q))\|\,.$$

Additive noise

Counterexample

- Expected value: (always)

$$\|\mathrm{gap}(\mathbb{E}\left[Q\right])\| \leqslant \|\mathbb{E}\left[\mathrm{gap}(Q)\right]\| \leqslant \|\mathrm{gap}(f)\| .$$

- Risk measures: (always)

$$\|\mathrm{gap}(\rho(Q))\| \leqslant \|\rho(\mathrm{gap}(Q))\| .$$

- Risk measures: (possible)

$$\|\mathrm{gap}(\rho(Q))\| \geqslant \|\mathrm{gap}(f)\| .$$

Additive noise

Counterexample

- Expected value: (always)

$$\|\text{gap}(\mathbb{E}[Q])\| \leqslant \|\mathbb{E}[\text{gap}(Q)]\| \leqslant \|\text{gap}(f)\|.$$

- Risk measures: (always)

$$\|\text{gap}(\rho(Q))\| \leqslant \|\rho(\text{gap}(Q))\|.$$

- Risk measures: (possible)

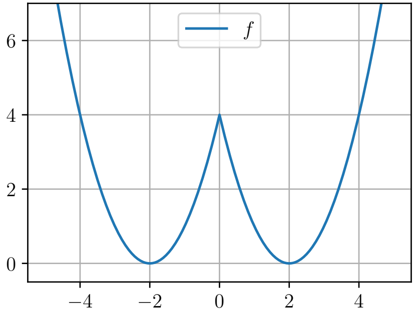
$$\|\text{gap}(\rho(Q))\| \geqslant \|\text{gap}(f)\|.$$

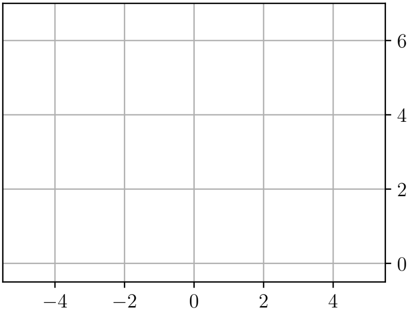
- Counterexample:

$$\begin{aligned} f(x) &= \min\left\{(x-2)^2,\, (x+2)^2\right\} \\ \rho(X) &= \sup_{\xi=\pm\frac{1}{2}} X \end{aligned}$$

Additive noise

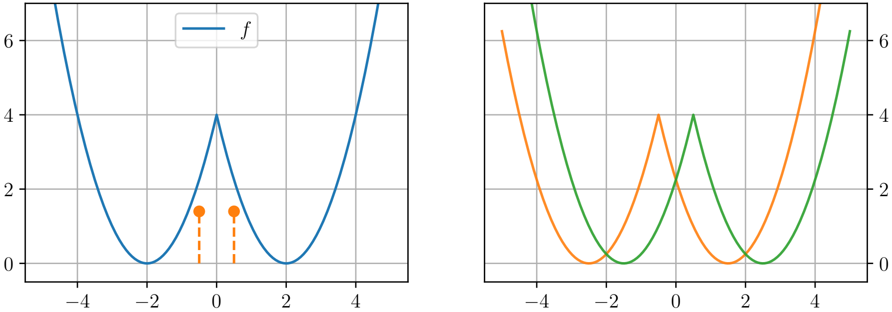
Counterexample





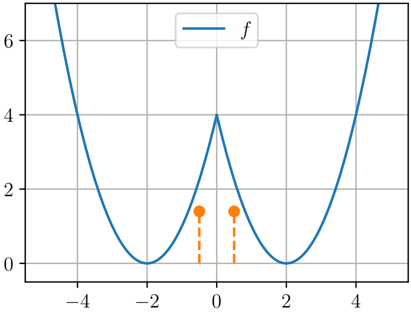
Additive noise

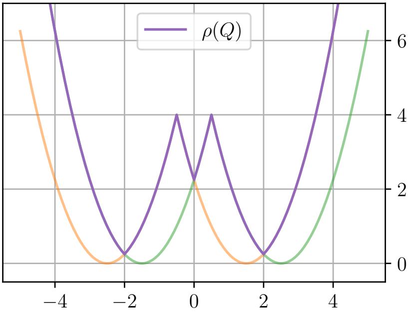
Counterexample



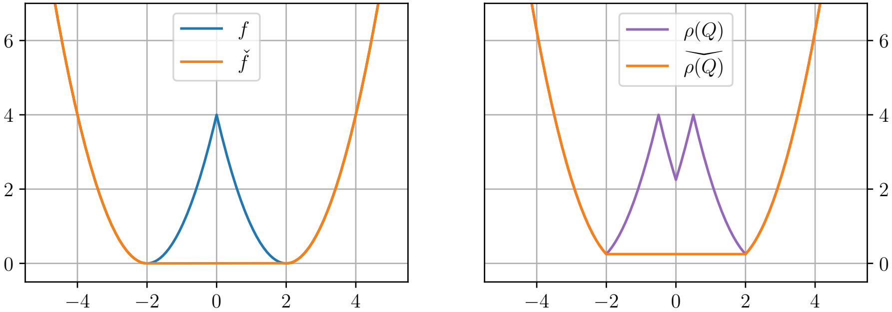
Additive noise

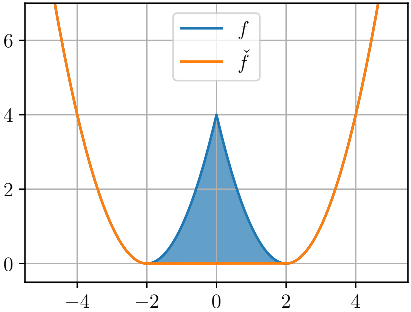
Counterexample

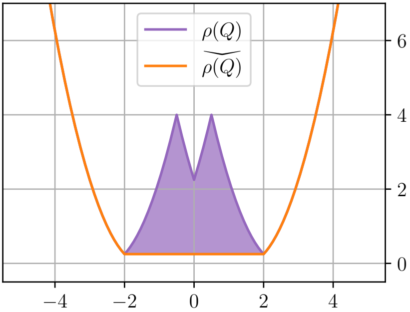




Additive noise
Counterexample

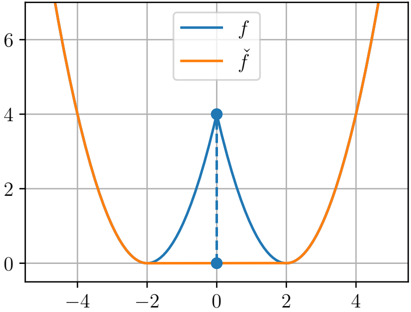


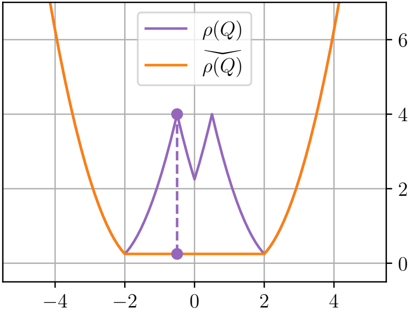




- Filled areas:

$$\|\text{gap}(f)\|_1 = \frac{16}{3}, \qquad \|\text{gap}(\rho(Q))\|_1 = \frac{22}{3}$$





- Filled areas and max heights:

$$\|\text{gap}(f)\|_1 = \frac{16}{3}, \qquad \|\text{gap}(\rho(Q))\|_1 = \frac{22}{3}$$

$$\|\text{gap}(f)\|_\infty = 4, \qquad \|\text{gap}(\rho(Q))\|_\infty = 3.75$$

Question

When the uncertainty is additive,

$$Q(x, \xi) = f(x - \xi) = \tau_\xi f(x),$$

the results regarding translation invariant norms still hold?

Answer

Not in general...

Question

When the uncertainty is additive,

$$Q(x, \xi) = f(x - \xi) = \tau_\xi f(x),$$

the results regarding translation invariant norms still hold?

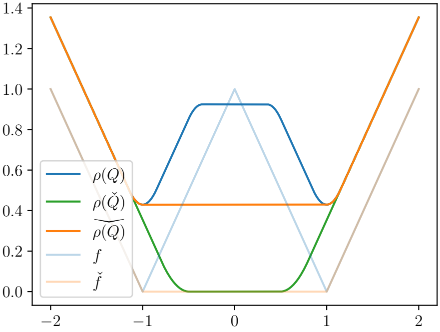
Answer

Not in general... **but yes for the uniform norm!**

Let $Q(x,\xi) = f(x-\xi) = \tau_\xi f(x)$.

- Always valid inequalities:

$$\begin{aligned}\rho(Q) - \widetilde{\rho(Q)} &\leqslant \rho(Q) - \rho(\check{Q}) \\ &\leqslant \rho(Q - \check{Q}).\end{aligned}$$



Additive Noise

Uniform Norm

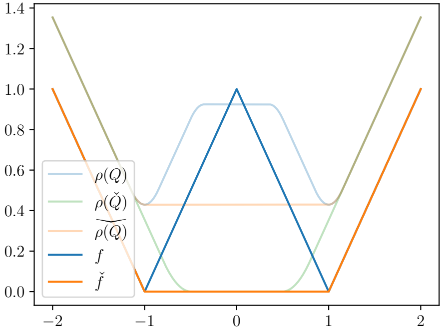
Let $Q(x, \xi) = f(x - \xi) = \tau_\xi f(x)$.

- Always valid inequalities:

$$\begin{aligned}\rho(Q) - \widetilde{\rho(Q)} &\leqslant \rho(Q) - \rho(\check{Q}) \\ &\leqslant \rho(Q - \check{Q}).\end{aligned}$$

- Uniform norm:

$$\begin{aligned}\|\rho(Q) - \widetilde{\rho(Q)}\|_\infty &\leqslant \|\rho(Q) - \rho(\check{Q})\|_\infty \\ &\leqslant \|\rho(Q - \check{Q})\|_\infty\end{aligned}$$



Additive Noise

Uniform Norm

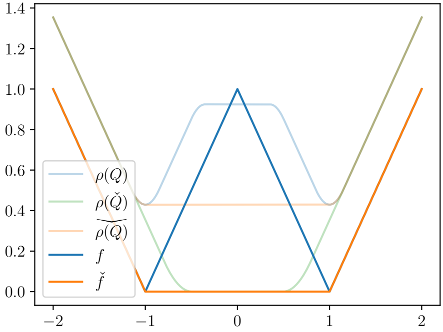
Let $Q(x, \xi) = f(x - \xi) = \tau_\xi f(x)$.

- Always valid inequalities:

$$\begin{aligned} \rho(Q) - \overbrace{\rho(Q)} &\leqslant \rho(Q) - \rho(\check{Q}) \\ &\leqslant \rho(Q - \check{Q}). \end{aligned}$$

- Uniform norm:

$$\begin{aligned} \|\rho(Q) - \overbrace{\rho(Q)}\|_\infty &\leqslant \|\rho(Q) - \rho(\check{Q})\|_\infty \\ &\leqslant \|\rho(Q - \check{Q})\|_\infty \\ &\leqslant \|f - \check{f}\|_\infty \end{aligned}$$



Additive Noise
Uniform Norm

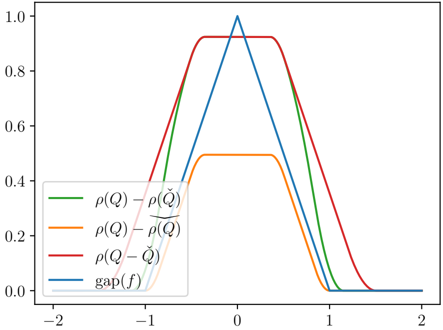
Let $Q(x,\xi) = f(x-\xi) = \tau_\xi f(x)$.

- Always valid inequalities:

$$\begin{aligned}\rho(Q) - \overbrace{\rho(Q)} &\leqslant \rho(Q) - \rho(\check{Q}) \\ &\leqslant \rho(Q - \check{Q}).\end{aligned}$$

- Uniform norm:

$$\begin{aligned}\|\rho(Q) - \overbrace{\rho(Q)}\|_\infty &\leqslant \|\rho(Q) - \rho(\check{Q})\|_\infty \\ &\leqslant \|\rho(Q - \check{Q})\|_\infty \\ &\leqslant \|f - \check{f}\|_\infty\end{aligned}$$



Application to cutting-planes algorithms

- Convex: approximations by cuts
- Non-convex: Construct better approximations from

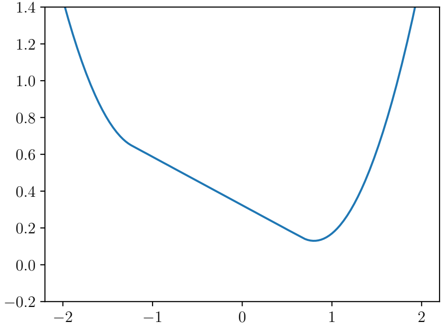
$$\rho(\check{Q}) \leqslant \widecheck{\rho(Q)} \leqslant \rho(Q).$$

- Integrate into Stochastic Dual Dynamic Programming.

Approximation by cuts

- Convex:

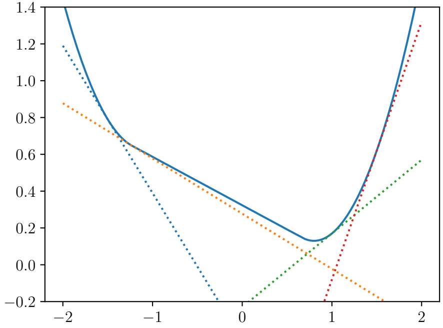
$$
\begin{aligned}
g(x) = \quad & \max_{a,b} \quad a^\top x - b \\
& \text{s.t.} \quad a^\top y - b \leqslant f(y),\ \forall y.
\end{aligned}
$$



Approximation by cuts

- Convex:

$$g(x) = \begin{array}{rl} \max\limits_{a,b} & a^\top x - b \\ \text{s.t.} & a^\top y - b \leqslant f(y),\ \forall y. \end{array}$$



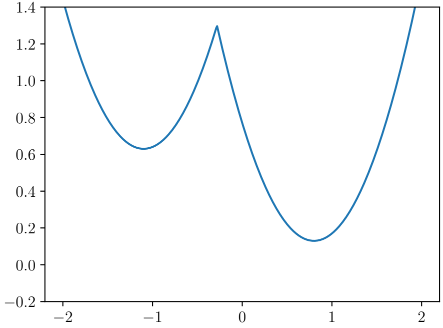
Approximation by cuts

- Convex:

$$g(x) = \max_{a,b} \quad a^\top x - b \qquad \text{s.t.} \quad a^\top y - b \leqslant f(y),\ \forall y.$$

- Non-convex:

$$\check{f}(x) = \max_{a,b} \quad a^\top x - b \qquad \text{s.t.} \quad a^\top y - b \leqslant f(y),\ \forall y.$$



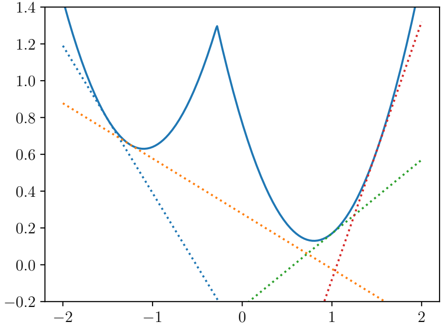
Approximation by cuts

- Convex:

$$g(x) = \max_{a,b} \; a^\top x - b \quad \text{s.t.} \quad a^\top y - b \leqslant f(y),\ \forall y.$$

- Non-convex:

$$\breve{f}(x) = \max_{a,b} \; a^\top x - b \quad \text{s.t.} \quad a^\top y - b \leqslant f(y),\ \forall y.$$



Only **convex relaxation** is approximable.

Approximation by cuts
Stochastic programs

What we want

Approximate $\mathbb{E}\left[Q\right]$ by cuts.

Approximation by cuts

Stochastic programs

What we want

Approximate $\mathbb{E}[Q]$ by cuts.

Standard method

- Calculate a cut for each scenario,
- Approximate via average cut.

Approximation by cuts

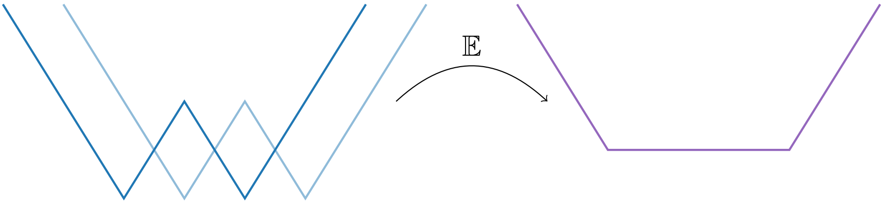
Stochastic programs

What we want

Approximate $\mathbb{E}\left[Q\right]$ by cuts.

Standard method

- Calculate a cut for each scenario,
- Approximate via average cut.



Approximation by cuts

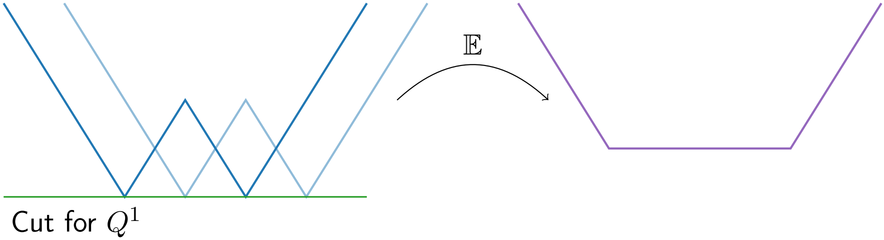
Stochastic programs

What we want

Approximate $\mathbb{E}[Q]$ by cuts.

Standard method

- Calculate a cut for each scenario,
- Approximate via average cut.

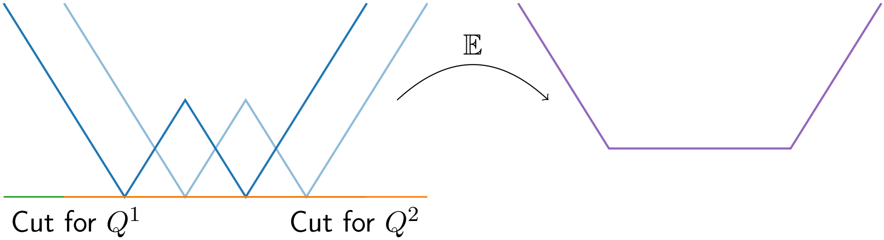


What we want

Approximate $\mathbb{E}\left[Q\right]$ by cuts.

Standard method

- Calculate a cut for each scenario,
- Approximate via average cut.



Approximation by cuts

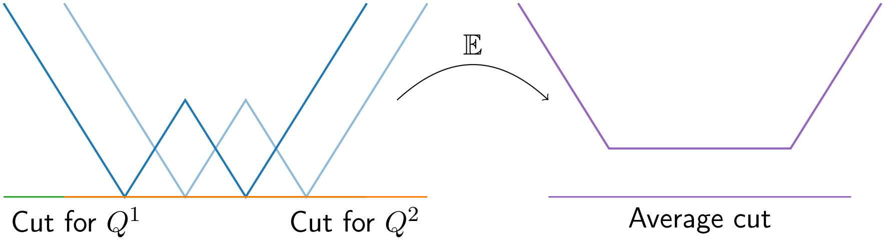
Stochastic programs

What we want

Approximate $\mathbb{E}[Q]$ by cuts.

Standard method

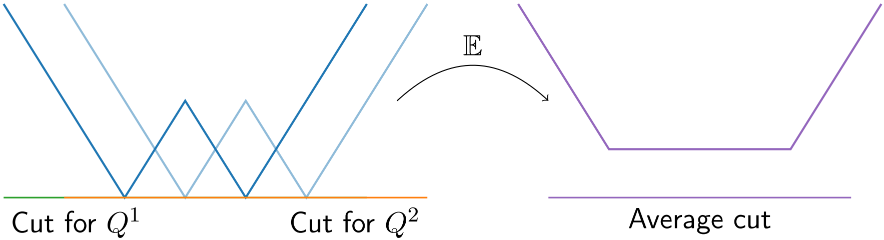
- Calculate a cut for each scenario,
- Approximate via average cut.



Approximation by cuts
Stochastic programs

Question

Why that cut was not tight?

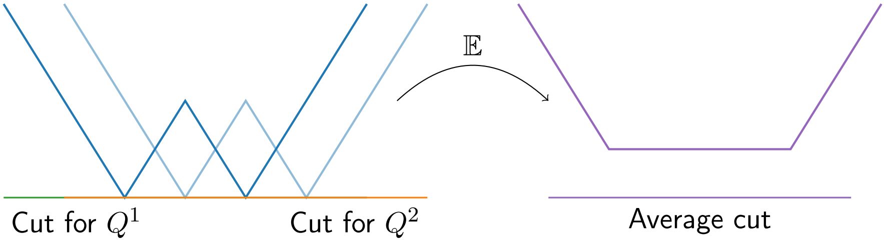


Question

Why that cut was not tight?

- Remember:

$$\mathbb{E}\,[\check{Q}] \leqslant \overgroup{\mathbb{E}\,[Q]} \leqslant \mathbb{E}\,[Q]\,.$$



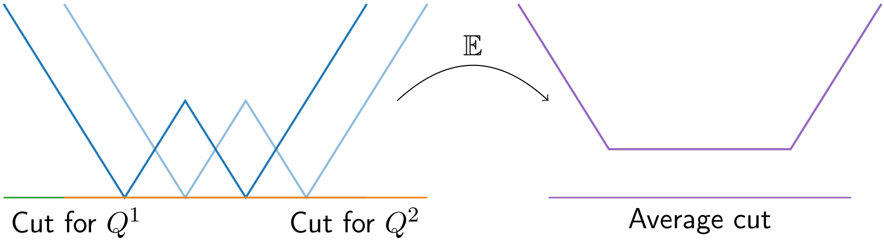
Question

Why that cut was not tight?

- Remember:

$$\mathbb{E}\,[\check{Q}] \leqslant \widetilde{\mathbb{E}\,[Q]} \leqslant \mathbb{E}\,[Q]\,.$$

- Cut for scenario ξ^i is only tight for $\check{Q}(\cdot,\xi^i)$.
- Average cut: only tight for $\mathbb{E}\,[\check{Q}]$.



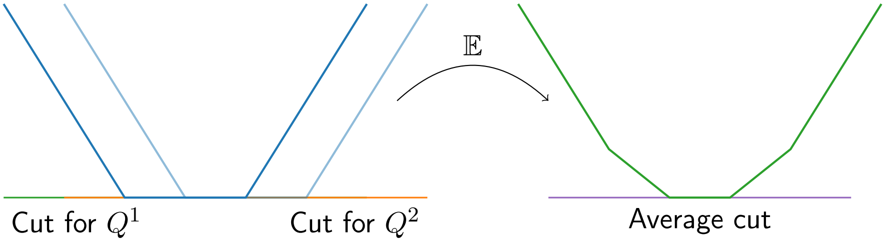
Question

Why that cut was not tight?

- Remember:

$$\mathbb{E}\,[\check{Q}] \leqslant \widetilde{\mathbb{E}\,[Q]} \leqslant \mathbb{E}\,[Q]\,.$$

- Cut for scenario ξ^i is only tight for $\check{Q}(\cdot, \xi^i)$.
- Average cut: only tight for $\mathbb{E}\,[\check{Q}]$.



Linked formulation

Question

How to directly approximate $\widetilde{\mathbb{E}\left[Q\right]}$ by cuts?

Cost-to-go function:

$$\begin{aligned} Q(x, \xi) = \quad \min_{y} \quad & c^\top x \\ \text{s.t.} \quad & Ax + By = b \\ & y \geqslant 0 \\ & y \in \mathbb{R}^n \times \mathbb{Z}^k. \end{aligned}$$

Linked formulation

Question

How to directly approximate $\widetilde{\mathbb{E}\left[Q\right]}$ by cuts?

Finite number of scenarios:

$$Q(x, \xi^i) = \begin{array}{rl} \min\limits_{y_i} & c_i^\top y_i \\ \text{s.t.} & A_i x + B_i y_i = b_i \\ & y_i \geqslant 0 \\ & y_i \in \mathbb{R}^n \times \mathbb{Z}^k. \end{array}$$

Linked formulation

Question

How to directly approximate $\widetilde{\mathbb{E}\left[Q\right]}$ by cuts?

Average over all scenarios:

$$\mathbb{E}\left[Q(x,\xi)\right] = \sum_{i=1}^{N} p_i \begin{array}{rl} \min\limits_{y_i} & c_i^\top y_i \\ \text{s.t.} & A_i x + B_i y_i = b_i \\ & y_i \geqslant 0 \\ & y_i \in \mathbb{R}^n \times \mathbb{Z}^k. \end{array}$$

Linked formulation

Question

How to directly approximate $\widetilde{\mathbb{E}[Q]}$ by cuts?

Build a linked program:

$$
\begin{aligned}
\mathbb{E}[Q(x,\xi)] = \min_{y_1,\ldots,y_N} \quad & \textstyle\sum_{i=1}^{N} p_i\, c_i^\top y_i \\
\text{s.t.} \quad & A_i x + B_i y_i = b_i, && \text{for } i = 1,\ldots,N \\
& y_i \geqslant 0, && \text{for } i = 1,\ldots,N \\
& y_i \in \mathbb{R}^n \times \mathbb{Z}^k, && \text{for } i = 1,\ldots,N.
\end{aligned}
$$

Linked formulation

Question

How to directly approximate $\widetilde{\mathbb{E}[Q]}$ by cuts?

Build a linked program:

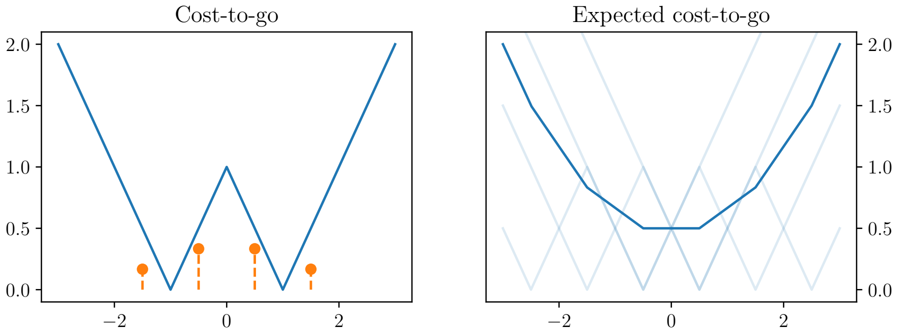
$$\begin{aligned} \mathbb{E}[Q(x,\xi)] = \min_{y_1,\dots,y_N} \quad & \textstyle\sum_{i=1}^N p_i\, c_i^\top y_i \\ \text{s.t.} \quad & A_i x + B_i y_i = b_i, && \text{for } i = 1,\dots,N \\ & y_i \geqslant 0, && \text{for } i = 1,\dots,N \\ & y_i \in \mathbb{R}^n \times \mathbb{Z}^k, && \text{for } i = 1,\dots,N. \end{aligned}$$

- Cuts for this problem can be tight for $\widetilde{\mathbb{E}[Q]}$!

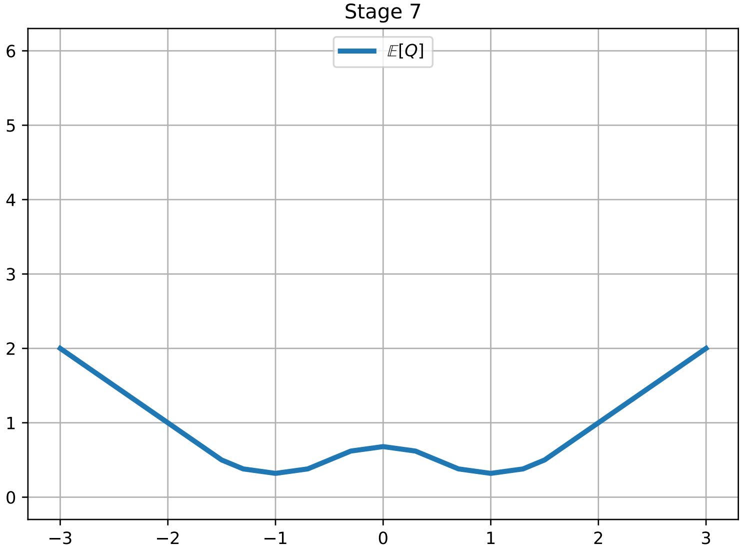
Multi-stage non-convex program,

$$\begin{aligned} Q_{t-1}(x_{t-1},\xi_t) = \quad \min_{x_t,u_t} \quad & |x_t| + \mathbb{E}\left[Q_t(x_t,\xi_{t+1})\right] \\ \text{s.t.} \quad & x_t = x_{t-1} + \xi_t + u_t \\ & u_t \in \{-1,1\}. \end{aligned}$$

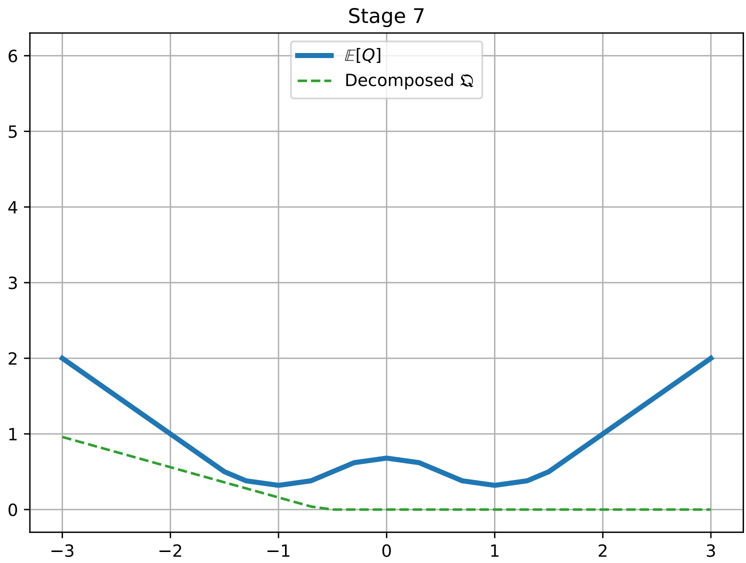
SDDP + strenghtened Benders cuts.



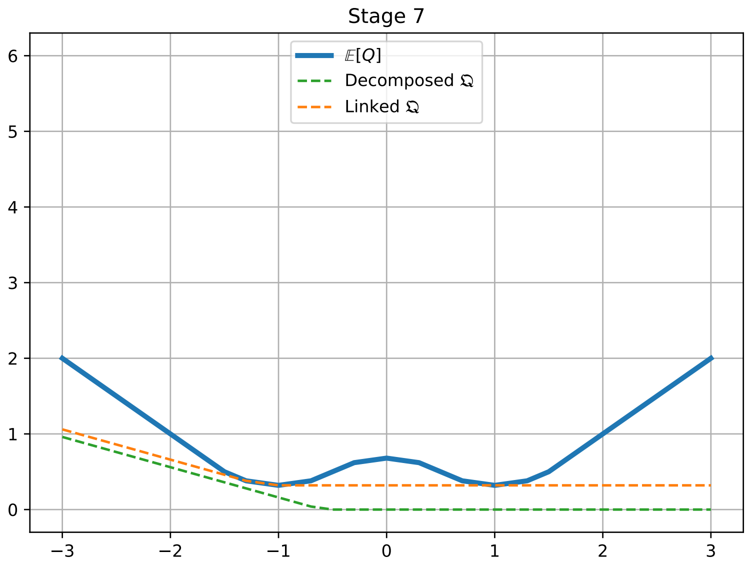
Computational experiments



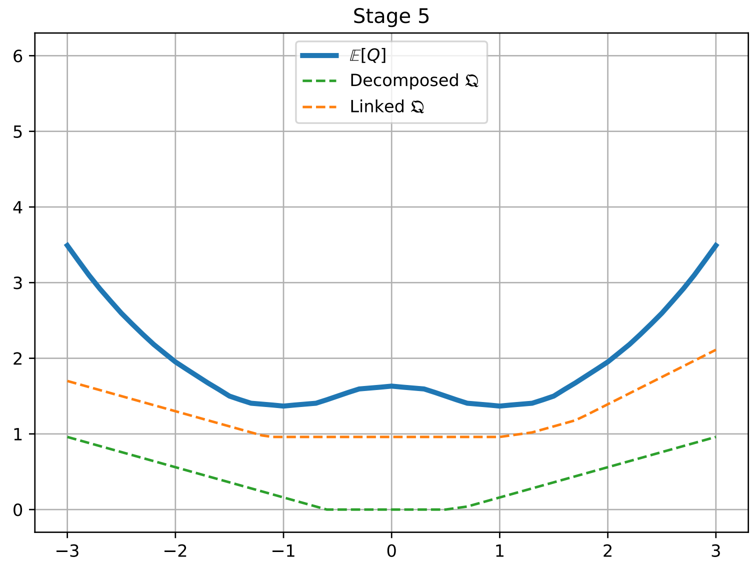
Computational experiments



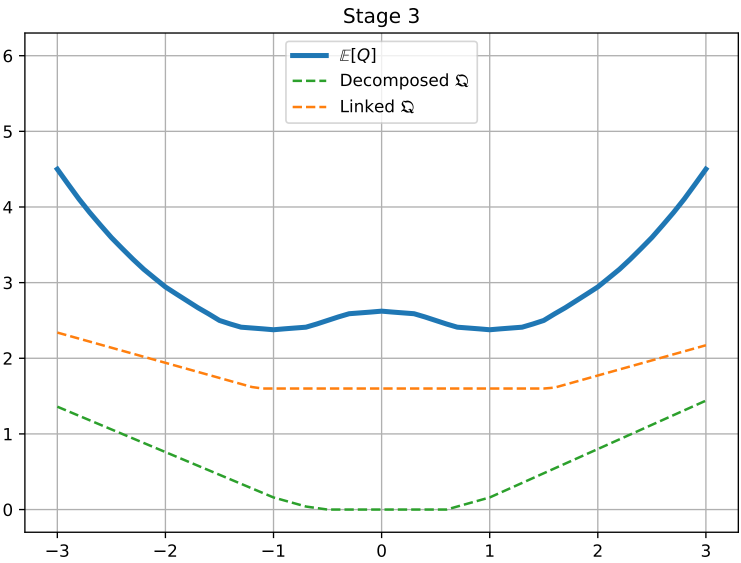
Computational experiments



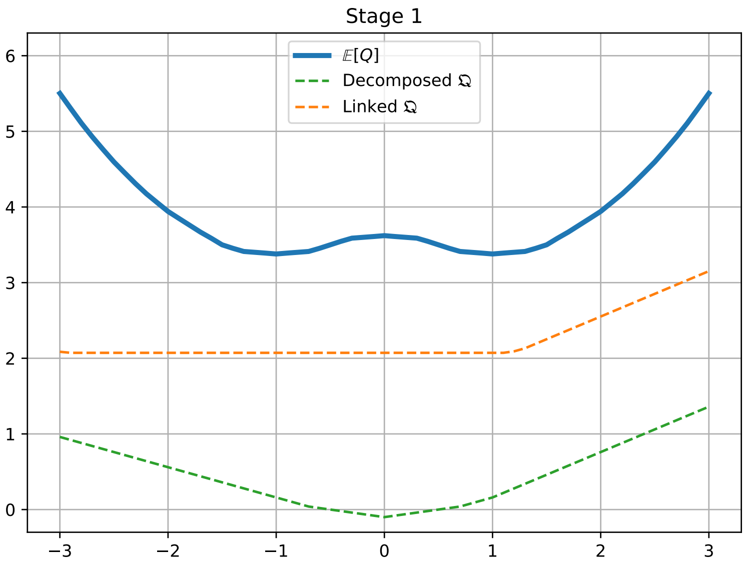
Computational experiments



Computational experiments



Computational experiments



Hydrothermal planning

$$
\begin{aligned}
Q_t(v_{t-1}, \xi_t) = \min_{v_t, q_t, s_t, g_t, df_t, f_t, z_t} \quad & c^\top g_t + g_{df}^\top df_t + \beta \bar{Q}_{t+1}(v_t) \\
\text{s.t.} \quad & v_t = v_{t-1} + \xi_t - q_t - s_t, \\
& q_t + M_I g_t + df_t + M_D f_t = d_t, \\
& 0 \leqslant v_t \leqslant \bar{v}, \quad 0 \leqslant q_t \leqslant \bar{q}, \quad 0 \leqslant s_t \\
& 0 \leqslant g_t \leqslant \bar{g}, \quad 0 \leqslant f_t \leqslant \bar{f}, \quad 0 \leqslant df_t, \\
& v_t \geqslant (1 - z_t)\, v_{\mathsf{MinOp}}, \\
& g_t \geqslant z_t\, g_0, \\
& z_t \in \{0, 1\}^n.
\end{aligned}
$$

- 500 iterations;
- 12 stages;
- $v_t \in \mathbb{R}^2$.

Table: Results for non-convex model with 2 subsystems.

	Cut types	
	Decomposed	Linked
Time (seconds)	759	16854
Iterations	500	500
Memory (GB)	0.418	1.089
Calculated cost (Bi R$)	10.410	11.170
Simulated cost (Bi R$)	12.227	12.209
Gap (%)	14.44	8.19

Table: Results for non-convex model with 2 subsystems.

	Cut types	
	Decomposed	Linked
Time (seconds)	759	16854
Iterations	500	500
Memory (GB)	0.418	1.089
Calculated cost (Bi R$)	10.410	11.170
Simulated cost (Bi R$)	12.227	12.209
Gap (%)	14.44	8.19

Table: Results for non-convex model with 2 subsystems.

	Cut types	
	Decomposed	Linked
Time (seconds)	759	16854
Iterations	500	500
Memory (GB)	0.418	1.089
Calculated cost (Bi R$)	10.410	11.170
Simulated cost (Bi R$)	12.227	12.209
Gap (%)	14.44	8.19

Table: Results for non-convex model with 2 subsystems.

	Cut types	
	Decomposed	Linked
Time (seconds)	18154	16854
Iterations	4000	500
Memory (GB)	0.418	1.089
Calculated cost (Bi R$)	10.410	11.170
Simulated cost (Bi R$)	12.227	12.209
Gap (%)	13.79	8.19

Table: Results for non-convex model with 2 subsystems.

	Cut types	
	Decomposed	Linked
Time (seconds)	18154	16854
Iterations	4000	500
Memory (GB)	0.418	1.089
Calculated cost (Bi R$)	10.410	11.170
Simulated cost (Bi R$)	12.227	12.209
Gap (%)	13.79	8.19

Table: Results for non-convex model with 2 subsystems.

	Cut types	
	Decomposed	Linked
Time (seconds)	18154	16854
Iterations	4000	500
Memory (GB)	0.418	1.089
Calculated cost (Bi R$)	10.410	11.170
Simulated cost (Bi R$)	12.227	12.209
Gap (%)	13.79	8.19

Future works

- Non-convexity measures: results for risk measures;
- Linked formulation: Large-scale problems;
- Linked formulation: linking only parts of the scenario tree.

Thank you!

Definition

The **conditional value-at-risk** is the coherent risk measure given by

$$\mathtt{CVaR}_\alpha[X] = \inf_{z \in \mathbb{R}} \; z + \frac{1}{1-\alpha}\mathbb{E}\big[[X - z]_+\big]$$

Linked formulation:

$$\begin{array}{rcll} \mathtt{CVaR}_\alpha[Q](x) = & \displaystyle\min_{z,t,y} & z + \frac{1}{1-\alpha}\sum_{i=1}^N p_i t_i & \\ & \text{s.t.} & c_i^\top y_i \leqslant t_i + z, & \text{for } i = 1, \dots, N \\ & & A_i x + B_i y_i = b_i, & \text{for } i = 1, \dots, N \\ & & y \geqslant 0,\; t \geqslant 0,\; z \in \mathbb{R}. & \end{array}$$