Séminaire de Mathématiques Appliquées du CERMICS

École des Ponts
ParisTech

Non-convexity measures

Bernardo Freitas Paulo da Costa
(Universidade Federal do Rio de Janeiro)

5 mars 2020

Non-convexity measures

Bernardo Freitas Paulo da Costa
with Iago Leal de Freitas, Filipe Goulart Cabral, Joari Paulo da Costa

Universidade Federal do Rio de Janeiro & ONS

5 Mars 2020

Motivation

Multistage, non-convex, stochastic optimization

- Costly calculation of Lagrangian cuts in SDDiP.
- Addition of binary variables to represent Lipschitz or Augmented Lagrangian cuts in SLDP makes iterations take more and more time.

Motivation

Multistage, non-convex, stochastic optimization

- Costly calculation of Lagrangian cuts in SDDiP.
- Addition of binary variables to represent Lipschitz or Augmented Lagrangian cuts in SLDP makes iterations take more and more time.

Table: A 4d, 8-stage, non-convex problem

	SB	SDDiP	SLDP
# cuts	100	100	250
LB	4.676	4.589	7.883
time (s)	7.3	4100	225

Motivation

Multistage, non-convex, stochastic optimization

- Costly calculation of Lagrangian cuts in SDDiP.
- Addition of binary variables to represent Lipschitz or Augmented Lagrangian cuts in SLDP makes iterations take more and more time.
- The complexity increase is not so dramatic in SDDiP, but still noticeable.

Table: A 4d, 8-stage, non-convex problem

	SB	SDDiP	SLDP
# cuts	100	100	250
LB	4.676	4.589	7.883
time (s)	7.3	4100	225
First iteration (s)	0.04	29	0.05
Last iteration (s)	0.10	150	1.7

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

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

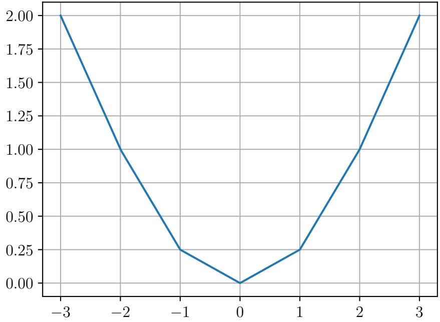


Figure: Example of cost-to-go.

Two-stage problems

Non-convex case

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

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

where the cost-to-go function is

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

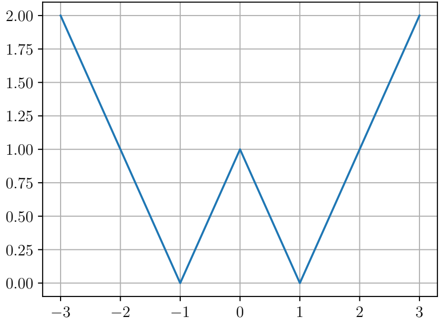


Figure: Example of cost-to-go.

Two-stage problems

Stochastic, non-convex case

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

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

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

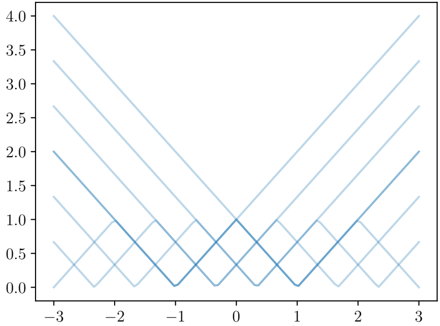


Figure: Random cost-to-go.

Two-stage problems

Stochastic, non-convex case

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

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

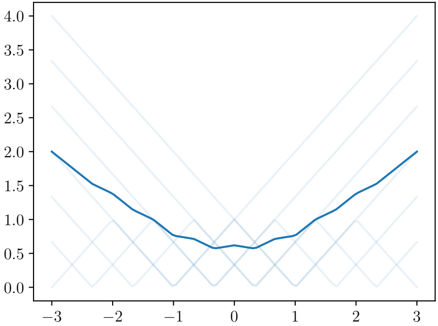
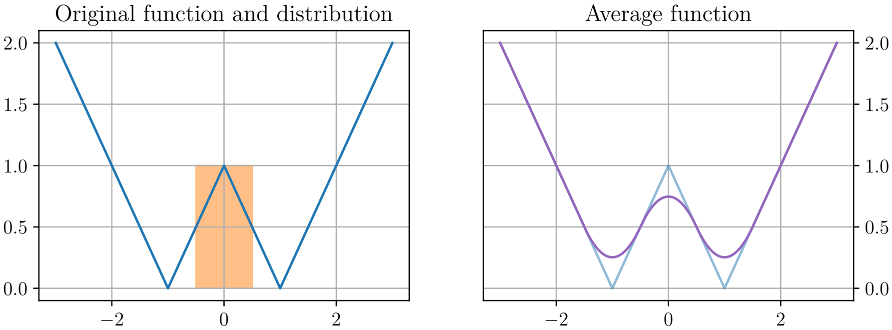


Figure: Expected cost-to-go.

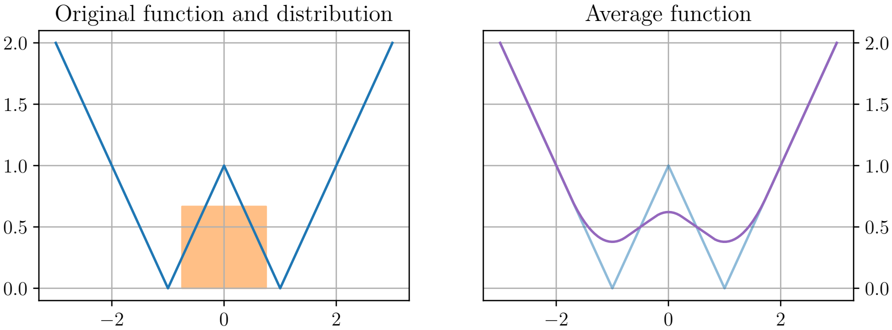
Non-convexity reduction

Uniform noises



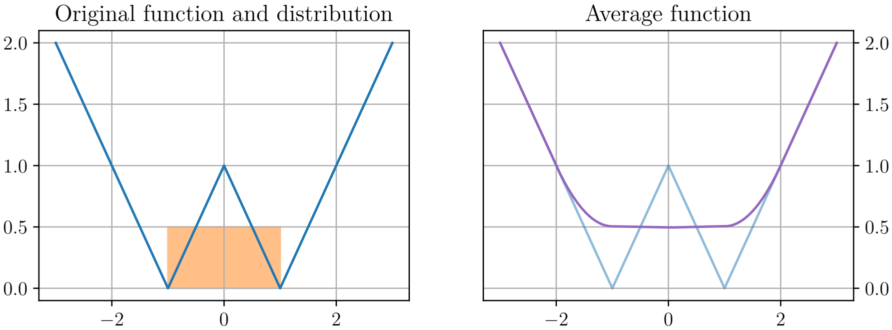
Non-convexity reduction

Uniform noises



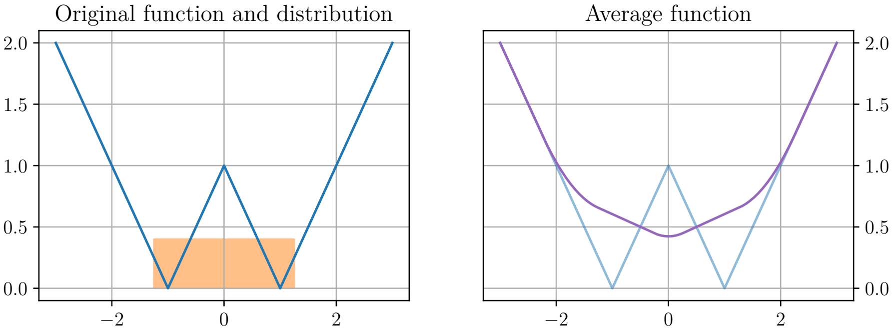
Non-convexity reduction

Uniform noises



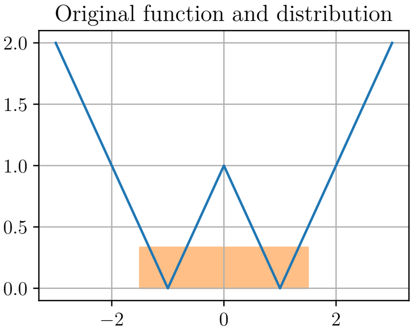
Non-convexity reduction

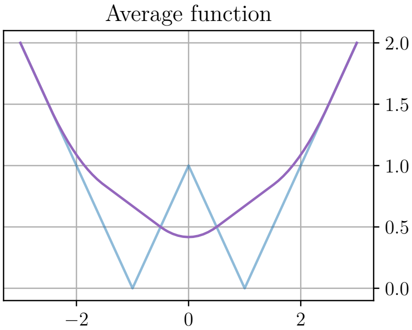
Uniform noises



Non-convexity reduction

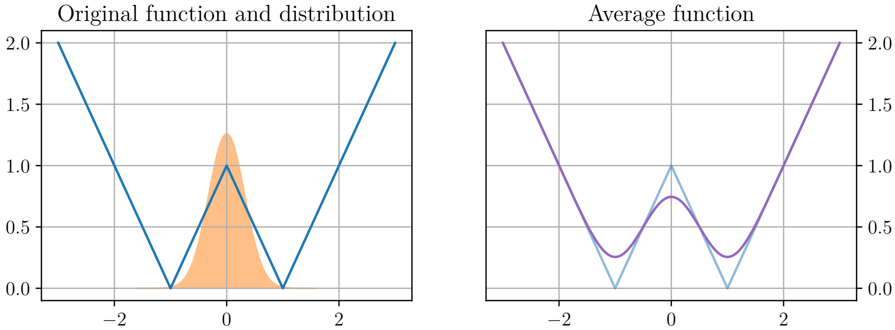
Uniform noises





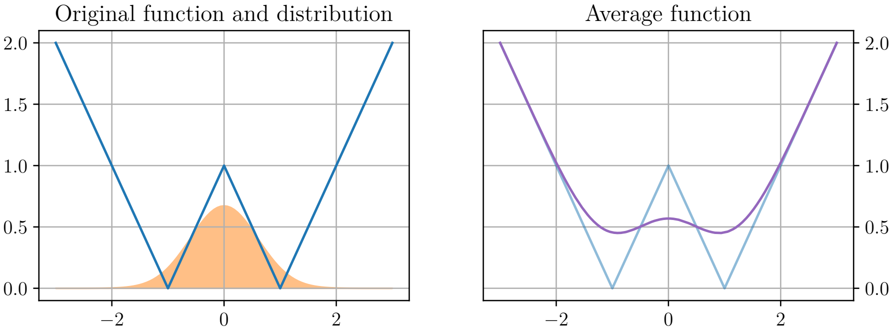
Non-convexity reduction

Gaussian noises

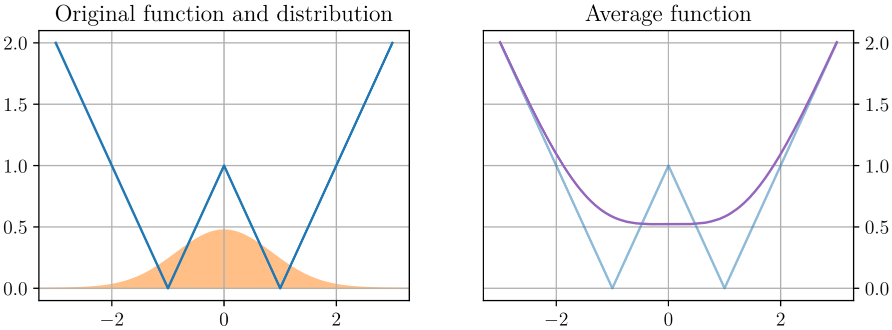


Non-convexity reduction

Gaussian noises

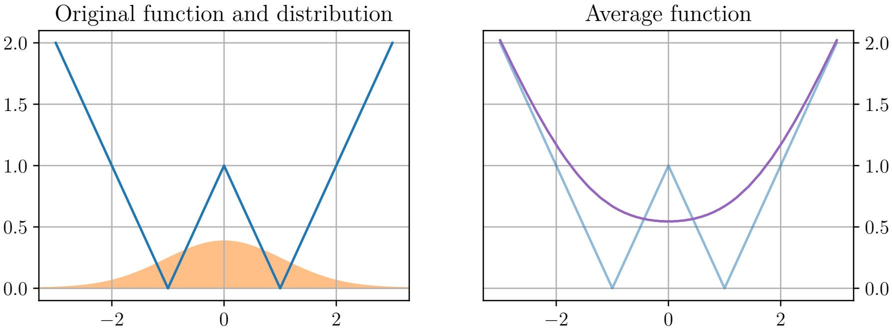


Non-convexity reduction
Gaussian noises



Non-convexity reduction

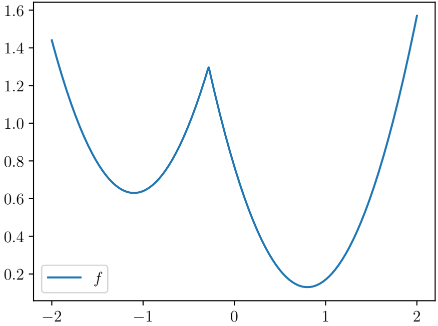
Gaussian noises



Measuring the non-convexity of a function

Convex relaxation

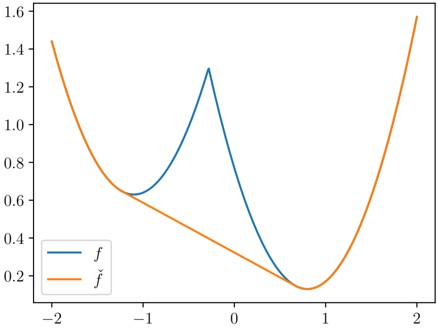
- Let f be a function.



Measuring the non-convexity of a function

Convex relaxation

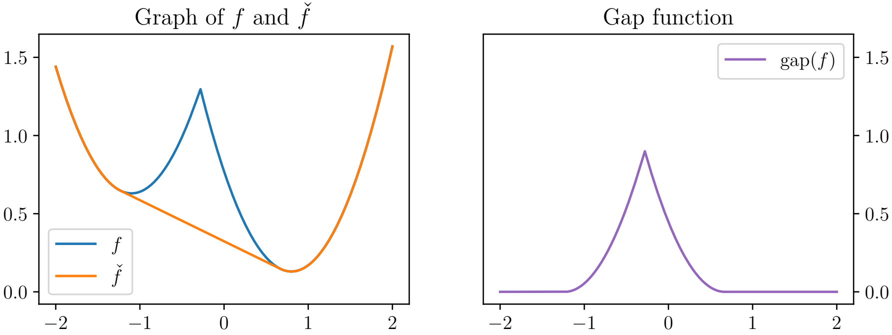
- Let f be a function.
- Its **convex relaxation** $\check{f}$ is the largest convex function everywhere less than f.



Measuring the non-convexity of a function

Gap function

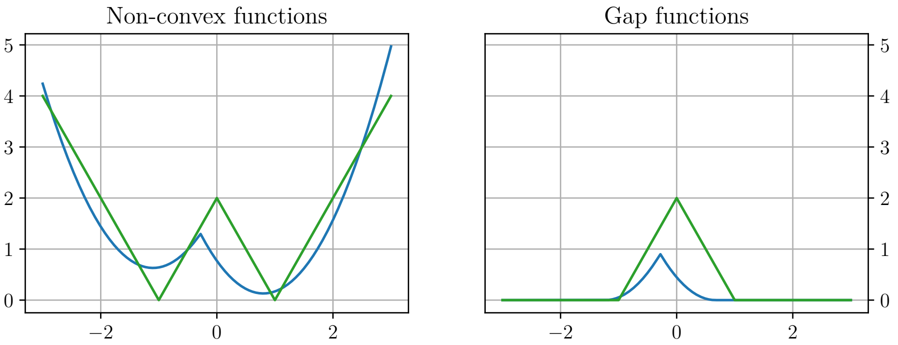
- $\mathrm{gap}(f) = f - \check{f}$;
- The gap is identically zero if and only if f is convex,
- $\mathrm{gap}(f)$ is always non-negative;



Measuring the non-convexity of a function
Gap comparisons

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

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



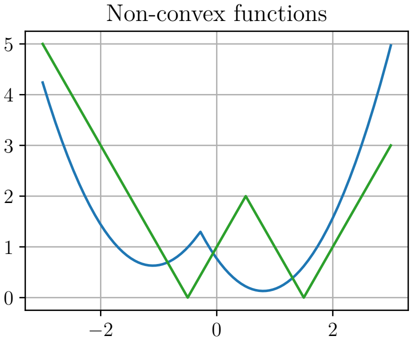
Measuring the non-convexity of a function

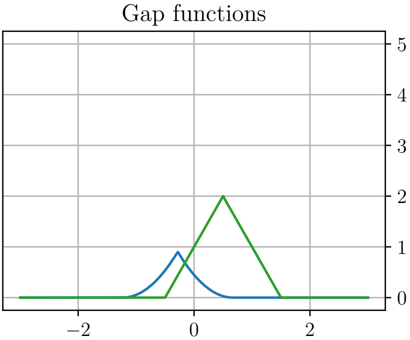
Gap comparisons

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

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

- But their gaps may not be comparable...





Measuring the non-convexity of a function

Monotone norms

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

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

Measuring the non-convexity of a function

Monotone norms

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

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

Definition

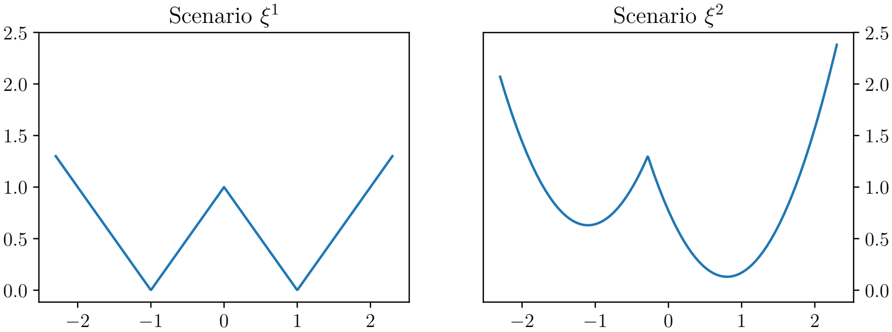
A function norm $\|\cdot\|$ is **monotone** if for all $g, h \geqslant 0$,

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

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

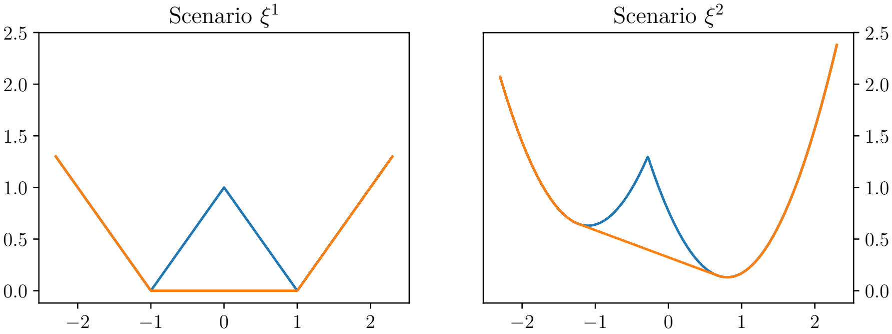
A bit of notation

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



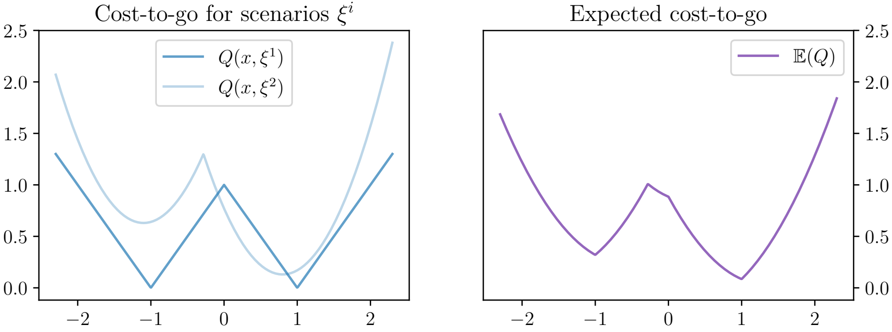
A bit of notation

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



A bit of notation

- Random variable: ξ.
- Random function: $Q(x, \xi)$.
- Convex relaxation $\check{Q}$: Defined for each realization $Q(\cdot, \xi)$.
- Average function: $\mathbb{E}\left[Q\right] = x \mapsto \mathbb{E}^{\xi}\left[Q(x, \xi)\right]$



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: $\check{Q} \leqslant Q$.

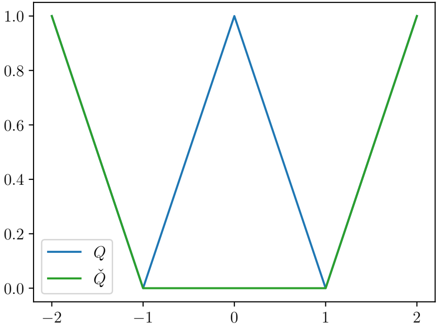


Figure: A realization of Q.

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:

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

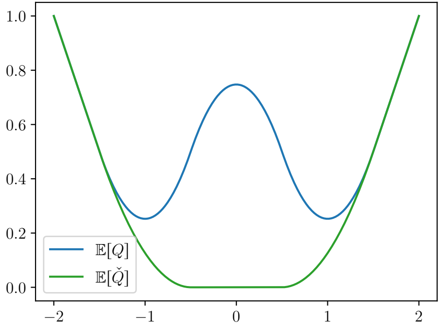


Figure: Average function.

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:

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

- And then:

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

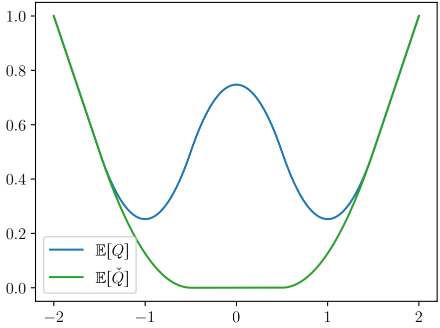


Figure: Average function.

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:

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

- And then:

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

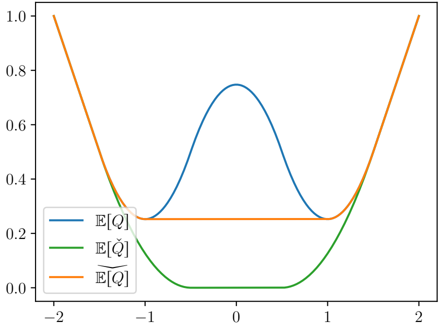


Figure: Average function.

Measuring the non-convexity of a function

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:

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

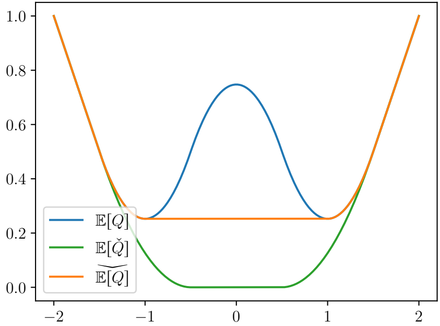


Figure: Average function.

Measuring the non-convexity of a function

Gap inequality

$\mathbb{E}\left[\check{Q}\right]$ and $\widecheck{\mathbb{E}\left[Q\right]}$: underapproximations to $\mathbb{E}\left[Q\right]$.

- $\mathbb{E}\left[\check{Q}\right] \leqslant \widecheck{\mathbb{E}\left[Q\right]} \leqslant \mathbb{E}\left[Q\right]$
- Rewriting:

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

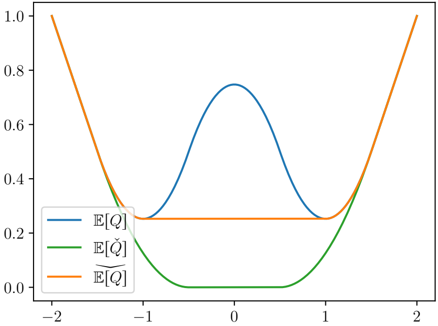


Figure: Average function.

Measuring the non-convexity of a function

Gap inequality

$\mathbb{E}\left[\check{Q}\right]$ and $\widecheck{\mathbb{E}\left[Q\right]}$: underapproximations to $\mathbb{E}\left[Q\right]$.

- $\mathbb{E}\left[\check{Q}\right] \leqslant \widecheck{\mathbb{E}\left[Q\right]} \leqslant \mathbb{E}\left[Q\right]$
- Rewriting:

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

- So:

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

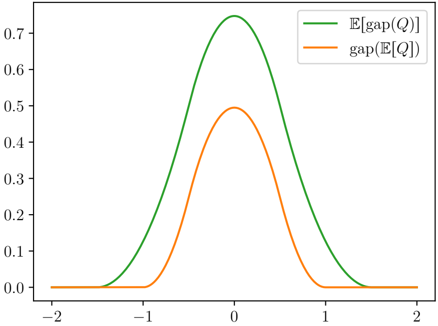


Figure: Gap functions.

Measuring the non-convexity of a function

Gap inequality

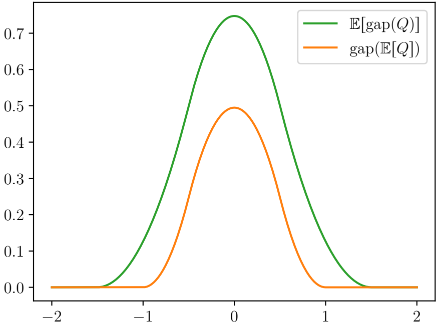
$\mathbb{E}[\check{Q}]$ and $\widehat{\mathbb{E}[Q]}$: underapproximations to $\mathbb{E}[Q]$.

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

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

- So:

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



- Monotone norm:

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

Back to optimization

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

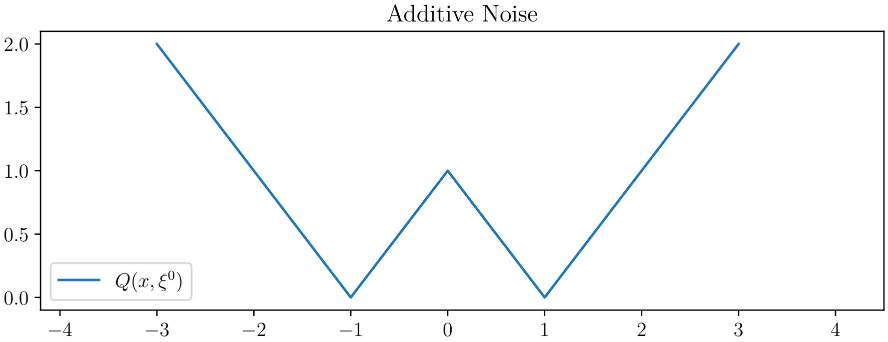
$$
\begin{aligned}
Q(x,\xi) = \min_{y} \quad & c^\top y \\
\text{s.t.} \quad & Ay = T(x-\xi) + b \\
& y \geqslant 0.
\end{aligned}
$$

Back to optimization

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

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

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

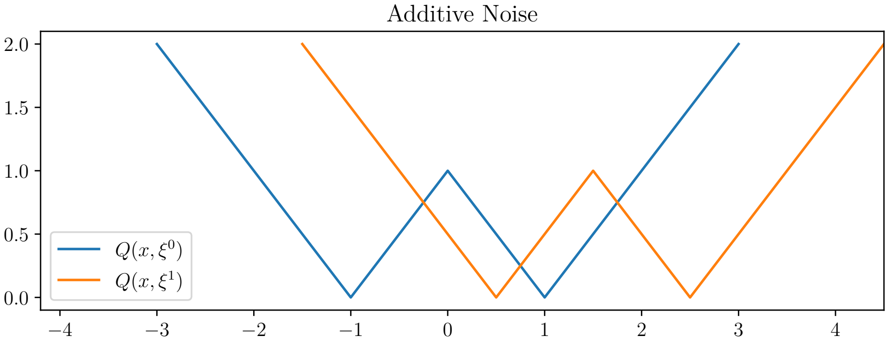


Back to optimization

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

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

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

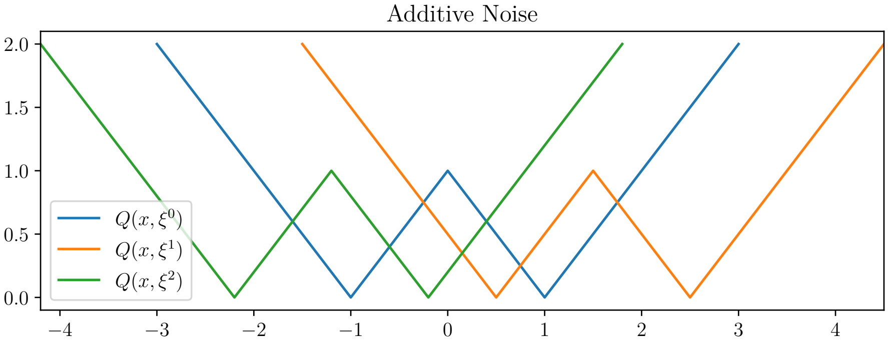


Back to optimization

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

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

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



Additive noise

- Let $Q(x, \xi) = f(x - \xi)$.
- The **translation operator** τ_a shifts a function by a,

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

Additive noise

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

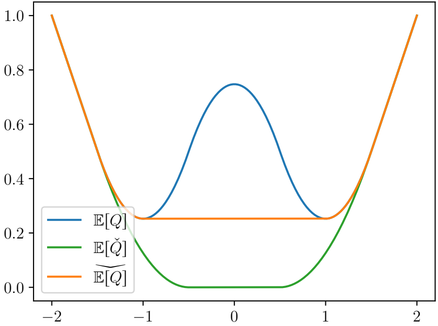
Additive noise

Inequality

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

- From the previous inequalities:

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



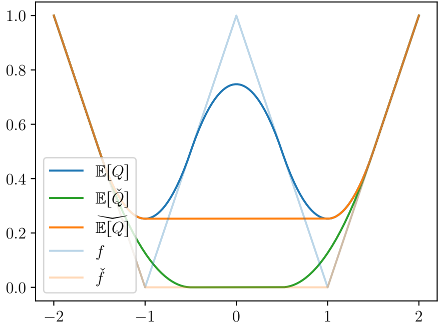
Additive noise

Inequality

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

- From the previous inequalities:

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



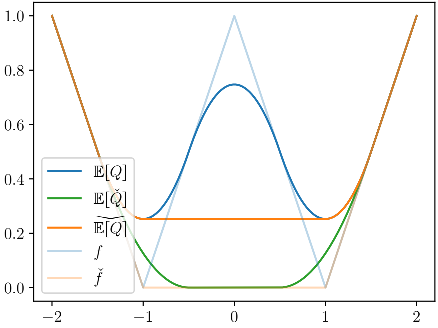
Additive noise

Inequality

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

- From the previous inequalities:

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



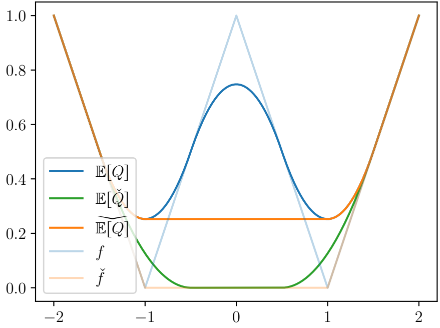
Additive noise

Inequality

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

- From the previous inequalities:

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



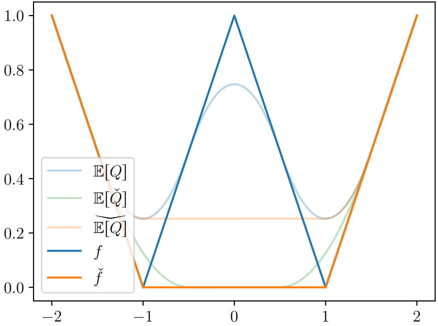
Additive noise

Inequality

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

- From the previous inequalities:

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



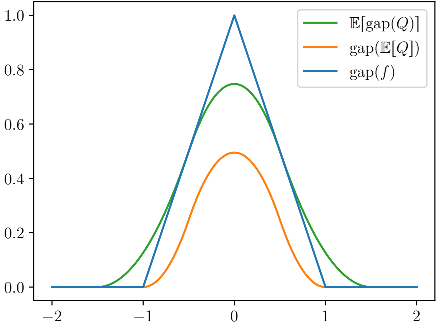
Additive noise

Inequality

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

- From the previous inequalities:

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



Additive noise

Uniform bounds

- Inequality for uniform norm:

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

Additive noise

Uniform bounds

- Inequality for uniform norm:

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

Theorem

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

where

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

Additive noise

Uniform bounds

- Inequality for uniform norm:

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

Theorem

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

where

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

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

Additive noise

Asymptotic behaviour

Theorem

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

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

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

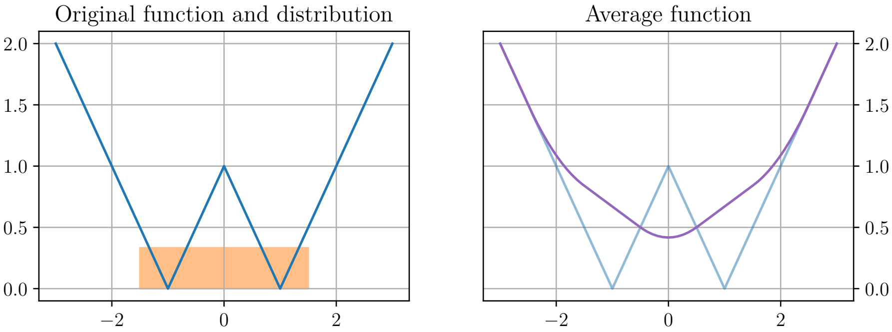
$$\|\mu_k\|_\infty \to 0,$$

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

Additive noise

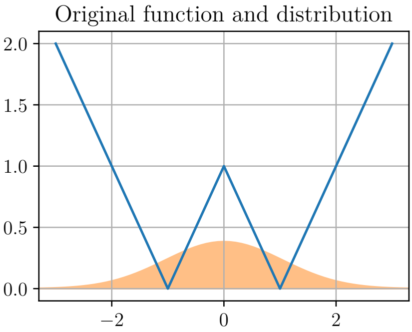
Asymptotic behaviour

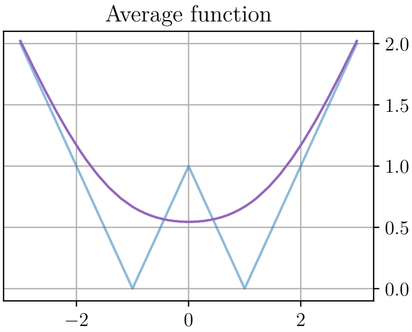
- $\xi_h \sim U[-h, h]$
- $\|\mu_h\|_\infty = \frac{1}{2h} \to 0$



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

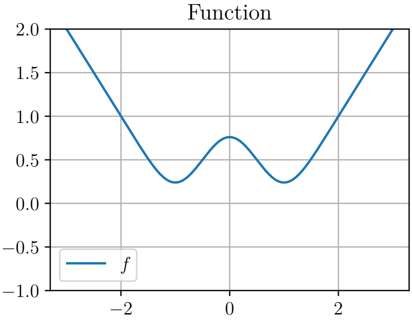
- Why the gap?
- For twice differentiable $f\colon (a,b) \to \mathbb{R}$,

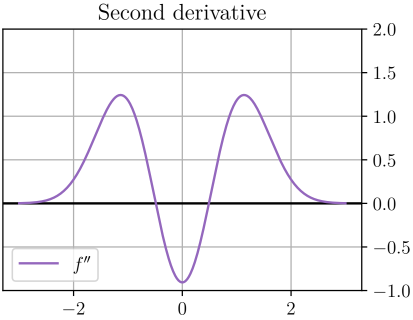
$$f \text{ is convex} \iff f'' \geqslant 0.$$

Another approach: second derivative

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

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

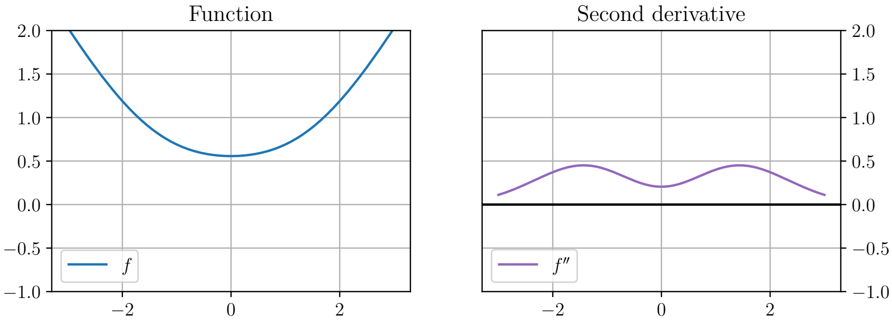




Another approach: second derivative

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

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



Another approach: second derivative

Idea

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

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

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

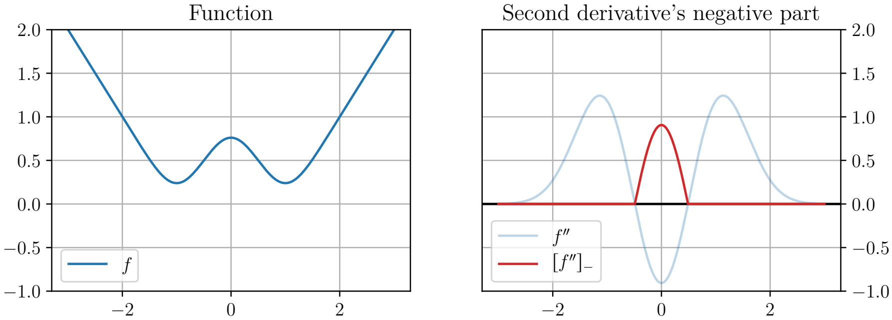
Another approach: second derivative

Idea

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

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

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



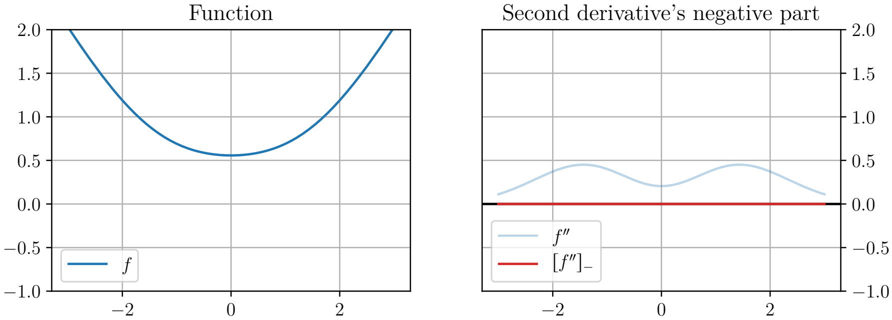
Another approach: second derivative

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.

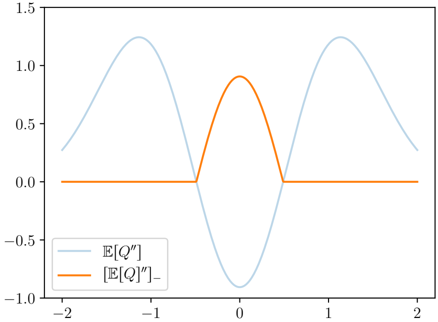
Another approach: second derivative

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

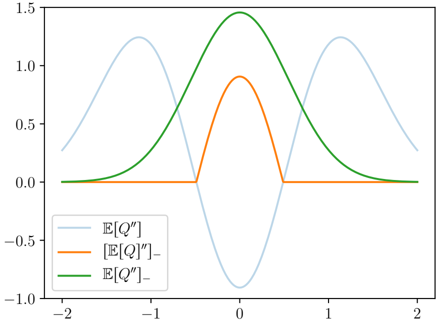
Let Q be a smooth random function,

- Exchanging expected value and derivative:

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

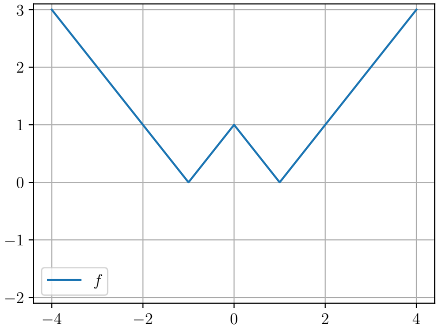
- And after some calculations:

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



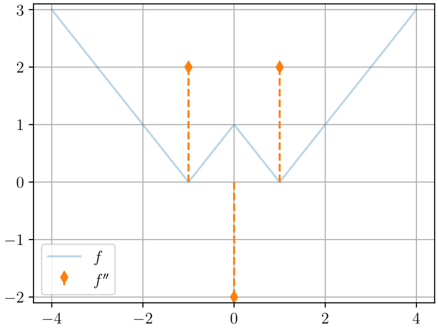
Derivative of cost-to-go

- The optimal value function of a mixed integer program is only **piecewise linear**.



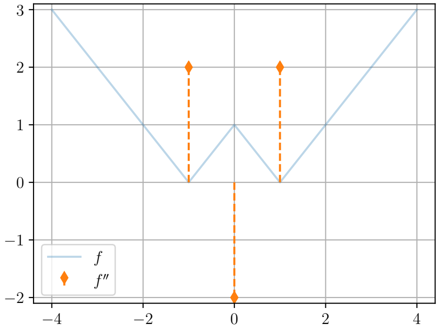
Derivative of cost-to-go

- The optimal value function of a mixed integer program is only **piecewise linear**.
- 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**.
- 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.

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

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

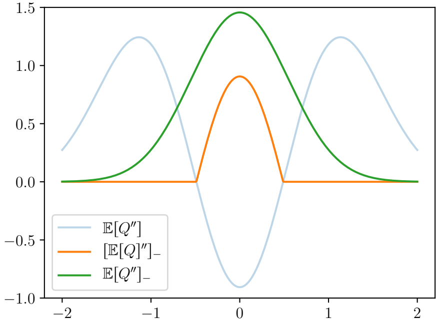
Another approach: second derivative

Main inequality

Let Q be a random continuous function.

- It still holds that:

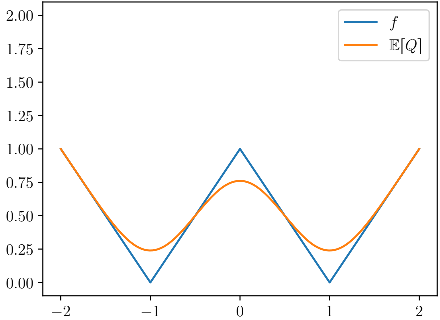
$$[\mathbb{E}[Q]'']_- \leqslant \mathbb{E}\big[[Q'']_-\big].$$



Reduction of non-convexity

Additive Noise

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

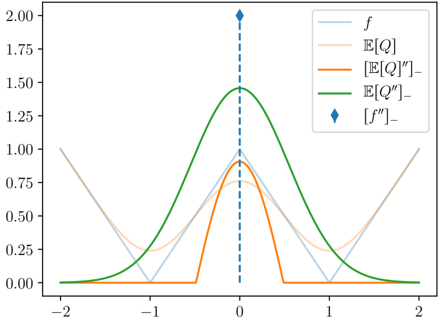


Reduction of non-convexity
Additive Noise

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

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

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



Reduction of non-convexity

Additive Noise

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

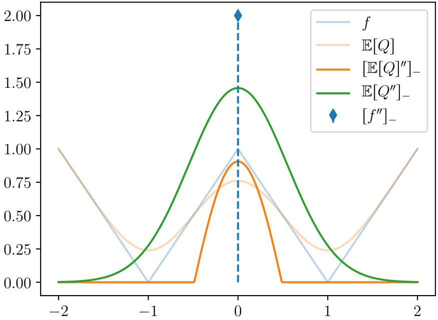
- f'' controls $\mathbb{E}\left[Q''\right]$:

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

- Uniform norm:

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

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



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 $[f'']_-$:

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

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

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

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

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

- Contains all rays:

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

- Is convex:

$$x,\, y \in K, \theta \in [0,1] \implies \theta x + (1-\theta) y \in K,$$

- Contains no lines:

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

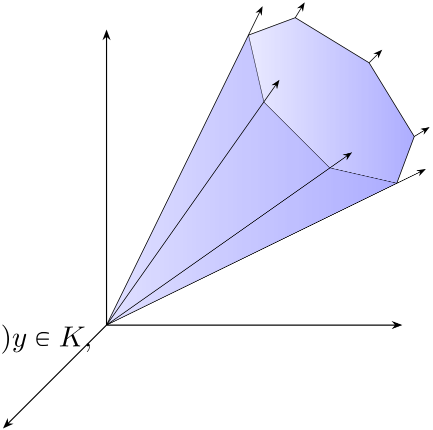


Figure: Example of convex cone.

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.

$\mathcal{M} = \mathrm{gap}$ 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:

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

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.

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 $D^2 f$,

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

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.

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.
- $\mathcal{R}(f) \geqslant 0$.

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

- Replace $\mathbb{E}$ in all formulas by a **risk measure**.

Risk-averse convexification

Coherent risk measures

- Replace $\mathbb{E}$ in all formulas by a **risk measure**.

Definition

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

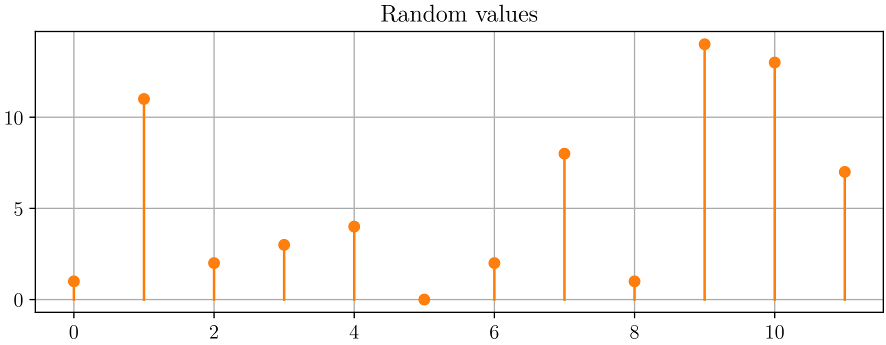
- Monotonicity: $X \leqslant Y \implies \rho(X) \leqslant \rho(Y)$;
- Translation equivariance: for all $a \in \mathbb{R}$, $\rho(X + a) = \rho(X) + a$;
- Convexity: for all $\lambda \in [0, 1]$,

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

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

Coherent risk measures

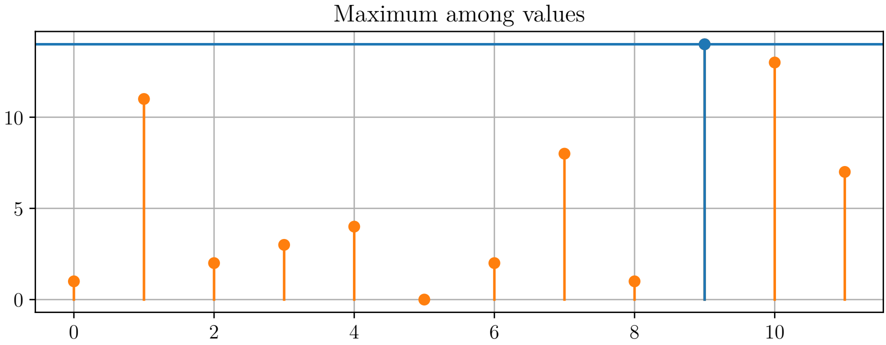
Examples of coherent risk measures:



Coherent risk measures

Examples of coherent risk measures:

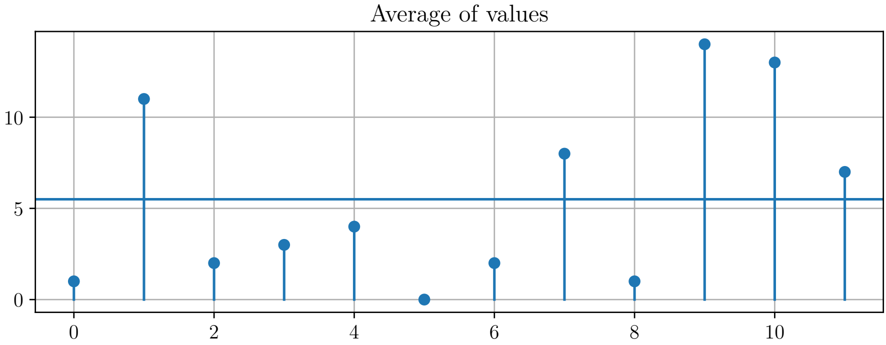
- Maximum value: $\sup X$;



Coherent risk measures

Examples of coherent risk measures:

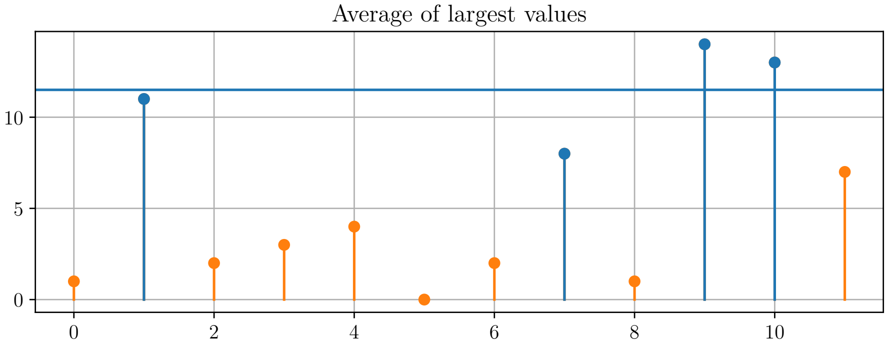
- Maximum value: $\sup X$;
- Expected value: $\mathbb{E}[X]$;



Coherent risk measures

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

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

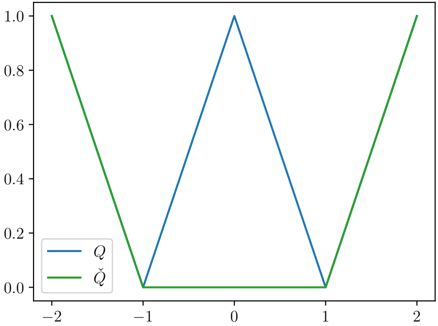
for a given family of probabilities $\mathcal{P}$.

Risk-averse convexification

Main inequality

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

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



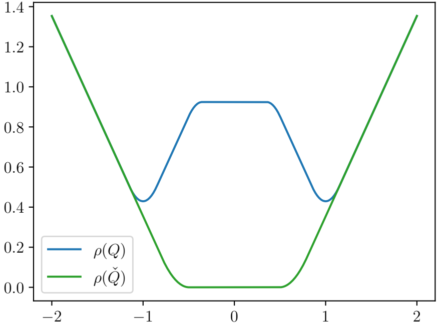
Risk-averse convexification

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

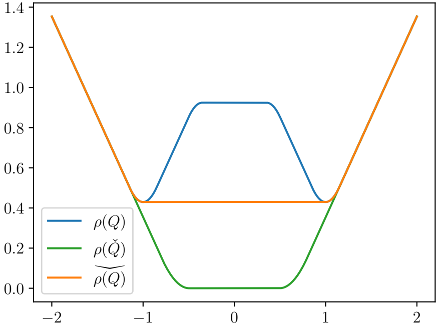
Let Q be a random function, ρ a coherent risk measure.

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

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

- Convex relaxation:

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

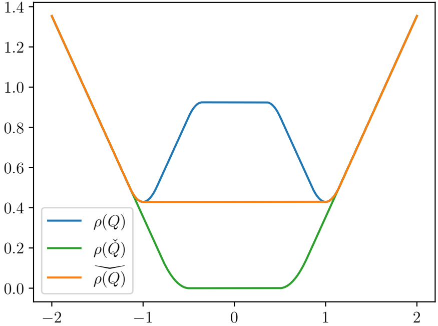


Risk-averse convexification

Gap inequality

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

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



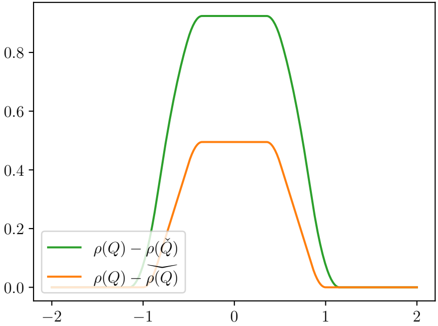
Risk-averse convexification

Gap inequality

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

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

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



Risk-averse convexification

Gap inequality

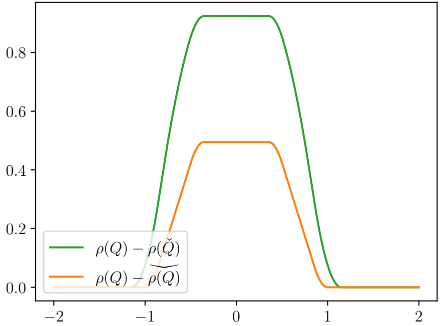
$\rho(\check{Q})$ and $\widetilde{\rho(Q)}$: underapproximations to $\rho(Q)$.

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

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

- 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

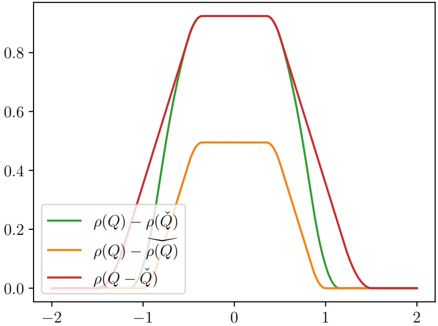
$\rho(\check{Q})$ and $\widetilde{\rho(Q)}$: underapproximations to $\rho(Q)$.

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

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

- 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,

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

do the results regarding translation invariant norms still hold?

Risk-averse convexification

Additive Noise

Question

When the uncertainty is additive,

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

do 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)

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

- Risk measures: (always)

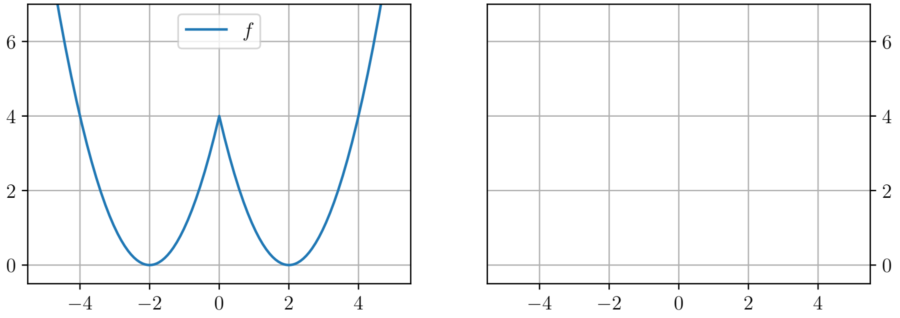
$$\|\mathrm{gap}(\rho(Q))\| \leqslant \|\rho(\mathrm{gap}(Q))\|\,.$$

- Risk measures: (possible)

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

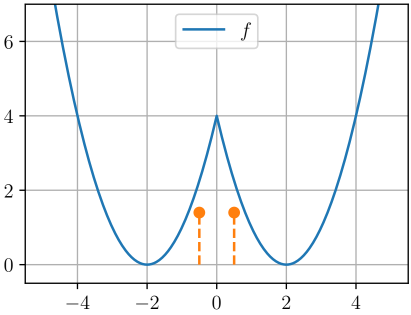
Additive noise

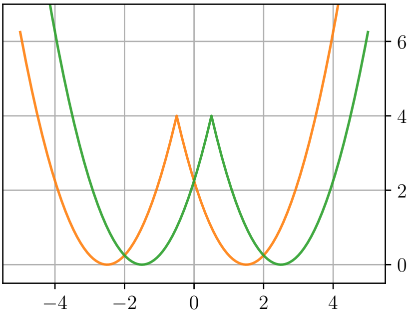
Counterexample



Additive noise

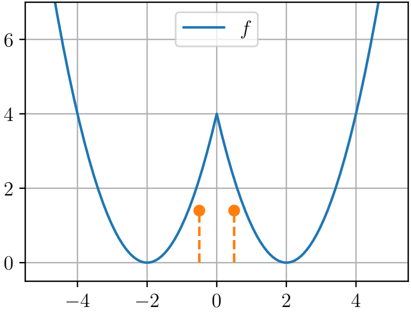
Counterexample

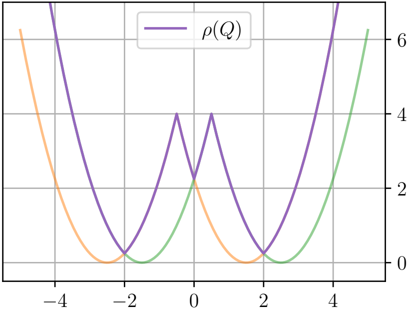




Additive noise

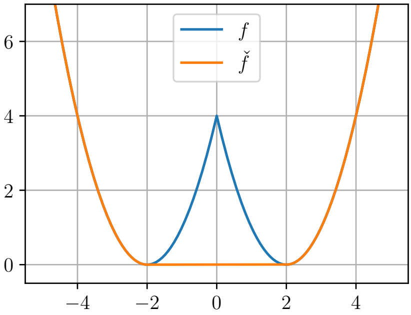
Counterexample

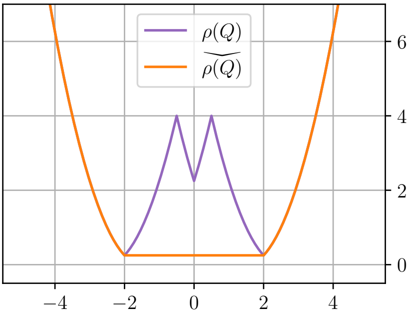




Additive noise

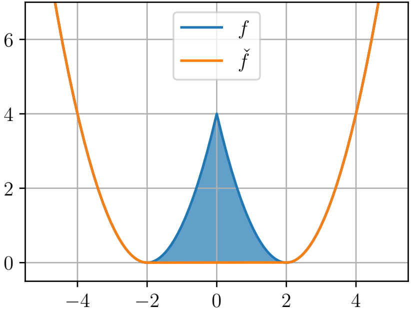
Counterexample

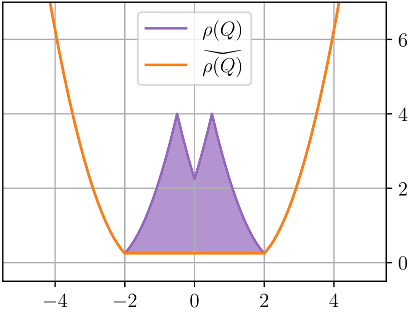




Additive noise

Counterexample



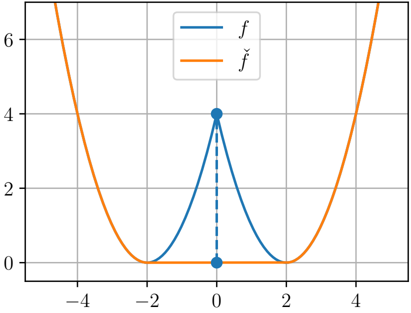


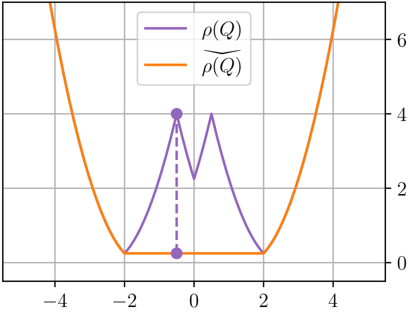
- Filled areas:

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

Additive noise

Counterexample





- Filled areas and max heights:

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

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

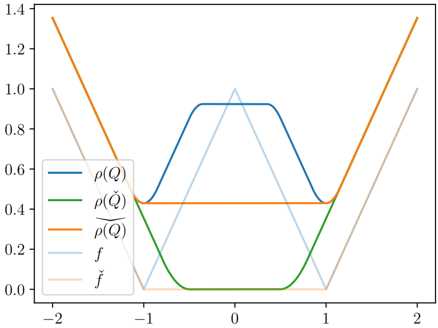
Additive Noise

Uniform Norm

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

- Always valid inequalities:

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



Additive Noise

Uniform Norm

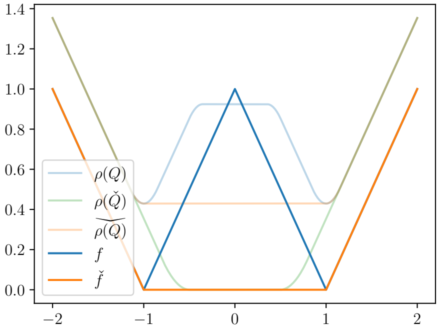
Let $Q(x,\xi) = f(x-\xi) = \tau_\xi f(x)$.

- Always valid inequalities:

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

- Uniform norm:

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



Additive Noise

Uniform Norm

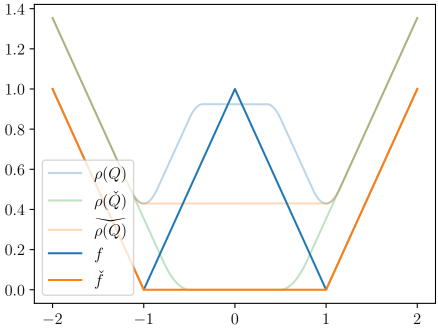
Let $Q(x,\xi) = f(x-\xi) = \tau_\xi f(x)$.

- 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:

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



Additive Noise

Uniform Norm

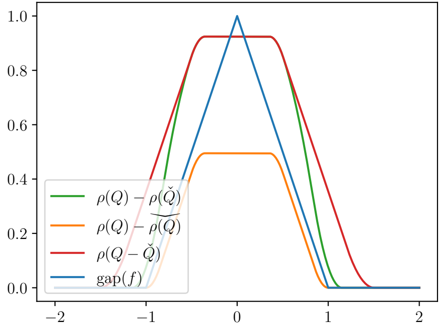
Let $Q(x,\xi) = f(x-\xi) = \tau_\xi f(x)$.

- Always valid inequalities:

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

- Uniform norm:

$$\|\rho(Q) - \widecheck{\rho(Q)}\|_\infty \leqslant \|\rho(Q) - \rho(\check{Q})\|_\infty \leqslant \|\rho(Q - \check{Q})\|_\infty \leqslant \|f - \check{f}\|_\infty$$



Approximation by cuts

Stochastic programs

What we want

Approximate $\mathbb{E}[Q]$ 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

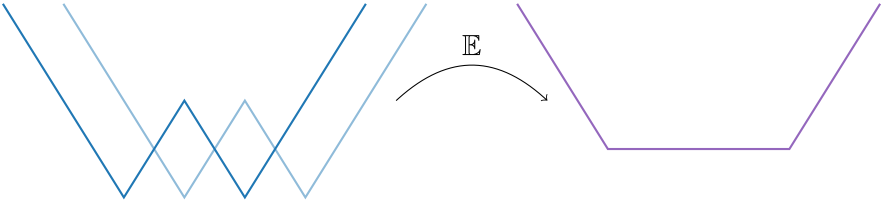
Stochastic programs

What we want

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

Standard method

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



Approximation by cuts

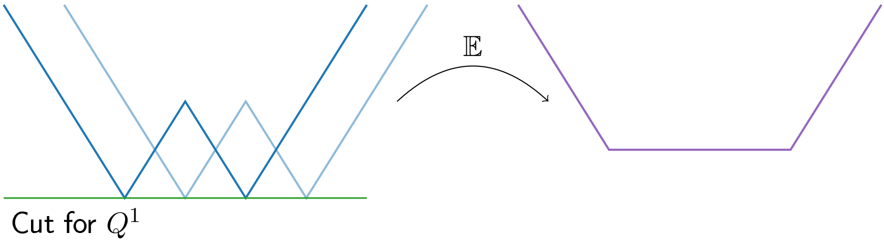
Stochastic programs

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

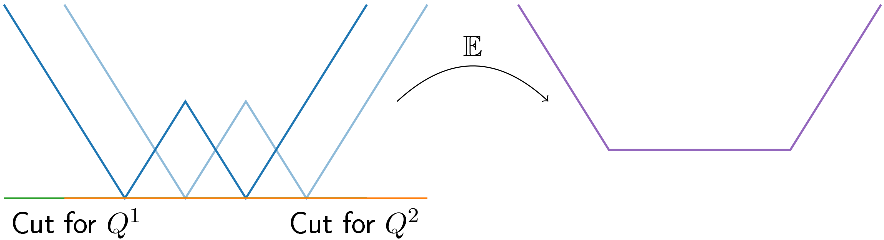
Stochastic programs

What we want

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

Standard method

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



Approximation by cuts

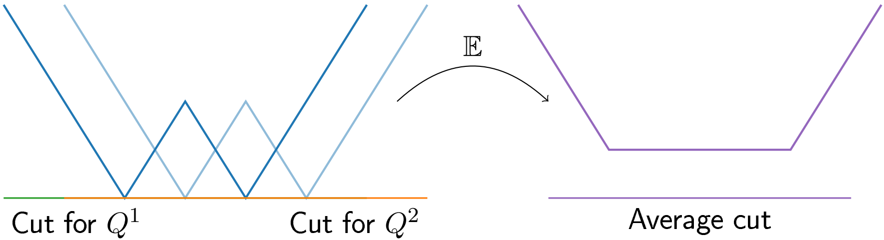
Stochastic programs

What we want

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

Standard method

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

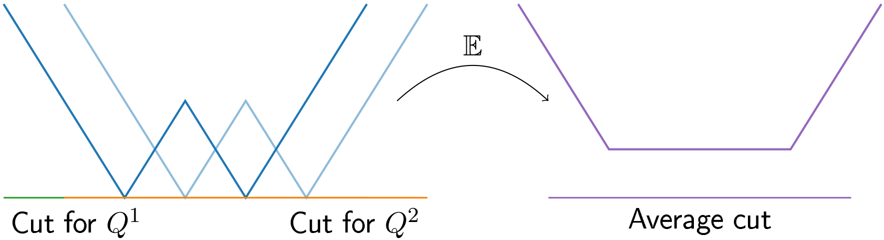


Approximation by cuts

Stochastic programs

Question

Why was that cut not tight?



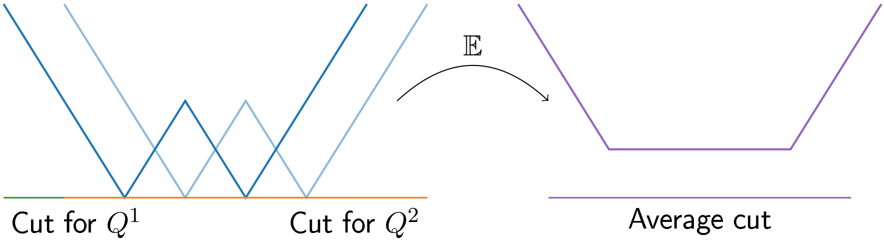
Approximation by cuts

Stochastic programs

Question

Why was that cut not tight?

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



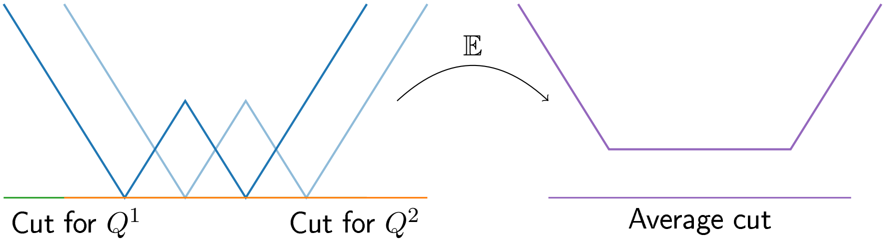
Approximation by cuts

Stochastic programs

Question

Why was that cut not tight?

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



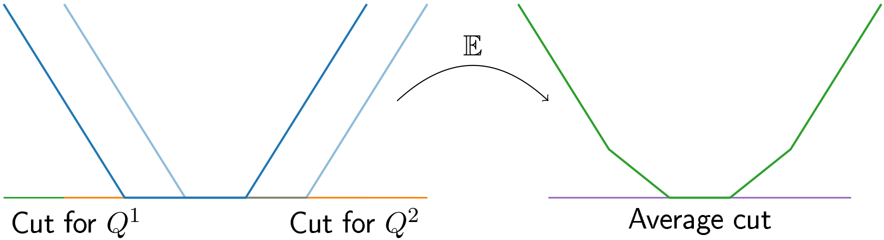
Approximation by cuts

Stochastic programs

Question

Why was that cut not tight?

- $\mathbb{E}\,[\check{Q}] \leqslant \widehat{\mathbb{E}\,[Q]} \leqslant \mathbb{E}\,[Q]\,.$
- 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 $\widetilde{\mathbb{E}[Q]}$ by cuts?

Finite number of scenarios:

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

Linked formulation

Question

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

Average over all scenarios:

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

Linked formulation

Question

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

Build a linked program:

$$
\begin{aligned}
\mathbb{E}\left[Q(x,\xi)\right] = \min_{\substack{y_1,\dots,y_N \\ x_1,\dots,x_N}} \quad & \textstyle\sum_{i=1}^{N} p_i\, c_i^\top y_i \\
\text{s.t.} \quad & A_i x_i + B_i y_i = b_i, && \text{for } i = 1,\dots,N \\
& y_i \geqslant 0, && \text{for } i = 1,\dots,N \\
& y_i \in \mathbb{R}^n \times \mathbb{Z}^k, && \text{for } i = 1,\dots,N. \\
& x_i = x && \text{for } i = 1,\dots,N
\end{aligned}
$$

Linked formulation

Question

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

Build a linked program:

$$
\begin{aligned}
\mathbb{E}\left[Q(x,\xi)\right] = \min_{\substack{y_1,\dots,y_N \\ x_1,\dots,x_N}} \quad & \textstyle\sum_{i=1}^{N} p_i\, c_i^\top y_i \\
\text{s.t.} \quad & A_i x_i + B_i y_i = b_i, && \text{for } i = 1,\dots,N \\
& y_i \geqslant 0, && \text{for } i = 1,\dots,N \\
& y_i \in \mathbb{R}^n \times \mathbb{Z}^k, && \text{for } i = 1,\dots,N. \\
& x_i = x && \text{for } i = 1,\dots,N
\end{aligned}
$$

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

Computational experiments

Table: A 4d, 8-stage, non-convex problem

	SB	SDDiP	SLDP	Relink
# cuts	100	100	250	250
LB	4.676	4.589	7.883	12.59
time (s)	7.3	4100	225	191
First iteration (s)	0.04	29	0.03	0.28
Last iteration (s)	0.10	150	1.7	0.67

Computational experiments

Table: A 4d, 8-stage, non-convex problem

	SB	SDDiP	SLDP	Relink
# cuts	100	100	250	250
LB	4.676	4.589	7.883	12.59
time (s)	7.3	4100	225	191
First iteration (s)	0.04	29	0.03	0.28
Last iteration (s)	0.10	150	1.7	0.67

“Relinking” is very expensive for several scenarios: this problem has only 10 scenario branches per stage.

Hydrothermal planning

Computational experiments

- 12 stages;
- 2 connected hydro reservoirs;
- Binary operational constraints.

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

Hydrothermal planning

Computational experiments

- 12 stages;
- 2 connected hydro reservoirs;
- Binary operational constraints.

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

Hydrothermal planning

Computational experiments

- 12 stages;
- 2 connected hydro reservoirs;
- Binary operational constraints.

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

Hydrothermal planning

Computational experiments

- 12 stages;
- 2 connected hydro reservoirs;
- Binary operational constraints.

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

Hydrothermal planning

Computational experiments

- 12 stages;
- 2 connected hydro reservoirs;
- Binary operational constraints.

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

Hydrothermal planning

Computational experiments

- 12 stages;
- 2 connected hydro reservoirs;
- Binary operational constraints.

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 work

- Large-scale problems: linking only parts of the scenario tree.
- Applications to SDDiP or SLDP.

Thank you!