# Gaining or Losing Perspective for Piecewise-Linear Under-Estimators of Convex Univariate Functions

Jon Lee[1], Daphne Skipper[2], Emily Speakman[3], **Luze Xu**[1]

[1]University of Michigan, Ann Arbor, MI

[2]U.S. Naval Academy, Annapolis, MD

[3]University of Colorado, Denver, CO

Sept 14, 2020

# Overview

1. Introduction
   - Background
   - Prior work
2. Piecewise-linear under-estimation and perspective
3. Analysis of convex power functions
   - quadratic function
   - non-quadratic function
   - monotonicity

# Overview


1 Introduction
- Background
- Prior work

2 Piecewise-linear under-estimation and perspective

3 Analysis of convex power functions
- quadratic function
- non-quadratic function
- monotonicity

# Background

We study the following model:

$$\begin{array}{ll} x = 0 & y = 0 \\ x \in [\ell, u] \ (0 \le \ell < u) & y = f(x) \end{array}$$

where $f$ is positive and convex on $[\ell, u]$.

# Background

We study the following model:

$$
\begin{array}{lll}
x = 0 & y = 0 & z = 0 \\
x \in [\ell, u] \ (0 \le \ell < u) & y = f(x) & z = 1
\end{array}
$$

where $f$ is positive and convex on $[\ell, u]$.

# Background

We study the following model:

$$\begin{array}{lll} x = 0 & y = 0 & z = 0 \\ x \in [\ell, u] \ (0 \leq \ell < u) & y = f(x) & z = 1 \end{array}$$

where $f$ is positive and convex on $[\ell, u]$.

$$\hat{D}_f(\ell, u) := \text{conv}\left( \{(0,0,0)\} \bigcup \left\{ (x, y, 1) \in \mathbb{R}^3 : \right.\right.$$
$$\left.\left. f(\ell) + \frac{f(u) - f(\ell)}{u - \ell}(x - \ell) \geq y \geq f(x),\ u \geq x \geq \ell \right\} . \right).$$

# Perspective relaxation

$$\hat{S}_f^*(\ell, u) := \text{cl}\Bigg\{(x, y, z) \in \mathbb{R}^3 \;\; : \\ \left(f(\ell) - \frac{f(u) - f(\ell)}{u - \ell}\ell\right) z + \frac{f(u) - f(\ell)}{u - \ell} x \geq y \geq z f(x/z), \\ uz \geq x \geq \ell z,\; 1 \geq z > 0,\; y \geq 0\Bigg\},$$

- $\hat{S}_f^*(\ell, u)$ is precisely the convex closure of $\hat{D}_f(\ell, u)$

# Perspective relaxation

$$\hat{S}_f^*(\ell,u) := \text{cl}\Bigg\{(x,y,z) \in \mathbb{R}^3 \ : \\ \left(f(\ell) - \frac{f(u)-f(\ell)}{u-\ell}\ell\right)z + \frac{f(u)-f(\ell)}{u-\ell}x \geq y \geq zf(x/z), \\ uz \geq x \geq \ell z,\ 1 \geq z > 0,\ y \geq 0\Bigg\},$$

- $\hat{S}_f^*(\ell,u)$ is precisely the convex closure of $\hat{D}_f(\ell,u)$
- Generally, working with $\hat{S}_f^*(\ell,u)$ implies using a cone solver (e.g., Mosek)
- Relaxations imply the possibility of using more general NLP solvers or even LP solvers

# Prior work

- The idea of using volume to compare relaxations was introduced by Lee and Morris [1994]
- Recently, Lee, Skipper, and Speakman [2019, 2020] applied the idea of using volumes to evaluate and compare the perspective relaxation with other relaxations of our disjunction

# Prior work

- The idea of using volume to compare relaxations was introduced by Lee and Morris [1994]
- Recently, Lee et al. [2019, 2020] applied the idea of using volumes to evaluate and compare the perspective relaxation with other relaxations of our disjunction
- Our focus is on relaxations related to natural piecewise-linear under-estimators of $f$

# Overview

1 Introduction
- Background
- Prior work

2 Piecewise-linear under-estimation and perspective

3 Analysis of convex power functions
- quadratic function
- non-quadratic function
- monotonicity

# Piecewise-linear under-estimation

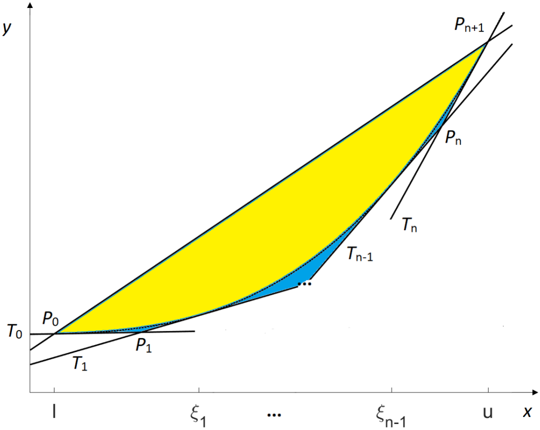



Figure: Piecewise-linear under-estimator

Linearization points: $\ell =: \xi_0 < \xi_1 < \cdots < \xi_{n-1} < \xi_n := u$

# Piecewise-linear under-estimation

The tangent line:

$$y = f(\xi_i) + f'(\xi_i)(x - \xi_i), \qquad (T_i)$$

Intersection point of $(T_i)$ and $(T_{i-1})$:

$$(x, y) := (\tau_i,\ f(\xi_i) + f'(\xi_i)(\tau_i - \xi_i)),\ \text{for } i = 1, \ldots, n, \qquad (P_i)$$

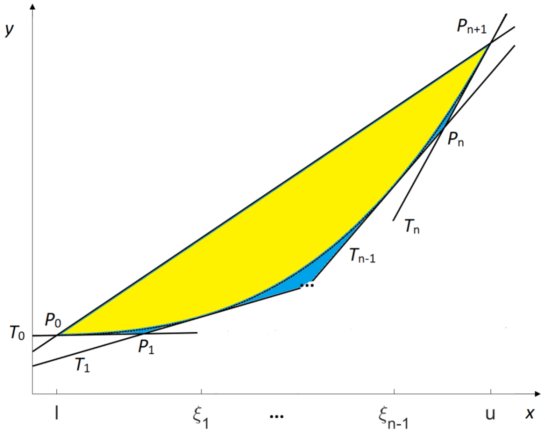
where

$$\tau_i := \frac{[f(\xi_i) - f'(\xi_i)\xi_i] - [f(\xi_{i-1}) - f'(\xi_{i-1})\xi_{i-1}]}{f'(\xi_{i-1}) - f'(\xi_i)}.$$

$$(x, y) := (\tau_0 := \ell, f(\ell)) \qquad (P_0)$$

$$(x, y) := (\tau_{n+1} := u, f(u)). \qquad (P_{n+1})$$

# Piecewise-linear under-estimation




The piecewise-linear under-estimator is referred to as $g$

# Volume of perspective relaxation

$$\hat{U}_f^*(\boldsymbol{\xi}) := \hat{S}_g^*(\ell, u) := \mathrm{cl}\Bigg\{(x, y, z) \in \mathbb{R}^3 \;:\; \left(f(\ell) - \frac{f(u) - f(\ell)}{u - \ell}\ell\right) z + \frac{f(u) - f(\ell)}{u - \ell} x \geq y \geq z g(x/z), \; uz \geq x \geq \ell z,\ 1 \geq z > 0,\ y \geq 0\Bigg\},$$

# Volume of perspective relaxation

$$\hat{U}^*_f(\xi) := \hat{S}^*_g(\ell, u) := \mathrm{cl}\Big\{(x, y, z) \in \mathbb{R}^3 \ :$$
$$\left(f(\ell) - \frac{f(u) - f(\ell)}{u - \ell}\ell\right) z + \frac{f(u) - f(\ell)}{u - \ell} x \geq y \geq zg(x/z),$$
$$uz \geq x \geq \ell z,\ 1 \geq z > 0,\ y \geq 0\Big\},$$

Theorem

$$\mathrm{vol}(\hat{U}^*_f(\xi)) = \frac{1}{6}\sum_{i=1}^{n}\left|\det\begin{pmatrix} \tau_0 & \tau_i & \tau_{i+1} \\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1}) \\ 1 & 1 & 1 \end{pmatrix}\right|.$$

# Volume of perspective relaxation

**Theorem**

$$\mathrm{vol}(\hat{U}_f^*(\xi)) = \frac{1}{6}\sum_{i=1}^{n} \left| \det \begin{pmatrix} \tau_0 & \tau_i & \tau_{i+1} \\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1}) \\ 1 & 1 & 1 \end{pmatrix} \right| .$$

- This set is a pyramid with apex $(x, y, z) = (0, 0, 0)$ and base equal to the intersection of $\hat{U}_f^*(\xi)$ with the hyperplane defined by the equation $z = 1$.

# Volume of perspective relaxation

**Theorem**

$$\mathrm{vol}(\hat{U}_f^*(\xi)) = \frac{1}{6}\sum_{i=1}^{n}\left|\det\begin{pmatrix} \tau_0 & \tau_i & \tau_{i+1} \\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1}) \\ 1 & 1 & 1 \end{pmatrix}\right|.$$

- This set is a pyramid with apex $(x, y, z) = (0, 0, 0)$ and base equal to the intersection of $\hat{U}_f^*(\xi)$ with the hyperplane defined by the equation $z = 1$.
- The height of the apex over the base is 1.

# Volume of perspective relaxation

**Theorem**

$$\mathrm{vol}(\hat{U}_f^*(\xi)) = \frac{1}{6}\sum_{i=1}^{n}\left|\det\begin{pmatrix} \tau_0 & \tau_i & \tau_{i+1} \\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1}) \\ 1 & 1 & 1 \end{pmatrix}\right|.$$

- This set is a pyramid with apex $(x, y, z) = (0, 0, 0)$ and base equal to the intersection of $\hat{U}_f^*(\xi)$ with the hyperplane defined by the equation $z = 1$.
- The height of the apex over the base is 1.
- The area of the base is computed by 2-d triangulation. Our triangles are $\mathrm{conv}\{P_0, P_i, P_{i+1}\}$, for $i = 1, \ldots, n$. The area of each triangle is $1/2$ of the absolute determinant of an appropriate $3 \times 3$ matrix.

# Volume of perspective relaxation

**Theorem**
$$\mathrm{vol}(\hat{U}_f^*(\xi)) = \frac{1}{6}\sum_{i=1}^{n}\left|\det\begin{pmatrix} \tau_0 & \tau_i & \tau_{i+1} \\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1}) \\ 1 & 1 & 1 \end{pmatrix}\right|.$$

- This set is a pyramid with apex $(x, y, z) = (0, 0, 0)$ and base equal to the intersection of $\hat{U}_f^*(\xi)$ with the hyperplane defined by the equation $z = 1$.
- The height of the apex over the base is 1.
- The area of the base is computed by 2-d triangulation. Our triangles are $\mathrm{conv}\{P_0, P_i, P_{i+1}\}$, for $i = 1, \ldots, n$. The area of each triangle is $1/2$ of the absolute determinant of an appropriate $3 \times 3$ matrix.
- So the volume of $\hat{U}_f^*(\xi)$ is simply the area of the base divided by 3. The formula follows.

# Volume of perspective relaxation

**Theorem**

$$\mathrm{vol}(\hat{U}_f^*(\xi)) = \frac{1}{6}\sum_{i=1}^{n}\left|\det\begin{pmatrix} \tau_0 & \tau_i & \tau_{i+1} \\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1}) \\ 1 & 1 & 1 \end{pmatrix}\right|.$$

**Corollary**

Assuming oracle access to $f$ and $f'$, we can compute $\mathrm{vol}(\hat{U}_f^*(\xi))$ in $\mathcal{O}(n)$ time.

With this volume, we can compare perspective relaxations of piecewise-linear under-estimators.

# Overview


1. Introduction
   - Background
   - Prior work
2. Piecewise-linear under-estimation and perspective
3. Analysis of convex power functions
   - quadratic function
   - non-quadratic function
   - monotonicity

# Analysis of convex power functions

Let $\hat{U}_p^*(\boldsymbol{\xi})$ denote $\hat{U}_f^*(\boldsymbol{\xi})$, with $f(x) := x^p$, $p > 1$.

# Analysis of convex power functions

Let $\hat{U}_p^*(\boldsymbol{\xi})$ denote $\hat{U}_f^*(\boldsymbol{\xi})$, with $f(x) := x^p$, $p > 1$.

### Theorem

Given $n \geq 2$, $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, we have that $\xi_i := \ell + \frac{i}{n}(u - \ell)$, for $i = 1, \ldots, n-1$, is the unique minimizer of $\text{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$, and the minimum volume is $\frac{1}{18}(u-\ell)^3 + \frac{(u-\ell)^3}{36n^2}$.

- Equally-spaced linearization points minimizes $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ when $p = 2$

# Analysis of quadratic function

### Theorem

Given $n \geq 2$, $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, we have that $\xi_i := \ell + \frac{i}{n}(u - \ell)$, for $i = 1, \ldots, n-1$, is the unique minimizer of $\text{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$, and the minimum volume is $\frac{1}{18}(u - \ell)^3 + \frac{(u-\ell)^3}{36n^2}$.

# Analysis of quadratic function

## Theorem

Given $n \geq 2$, $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, we have that $\xi_i := \ell + \frac{i}{n}(u - \ell)$, for $i = 1, \ldots, n-1$, is the unique minimizer of $\mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$, and the minimum volume is $\frac{1}{18}(u-\ell)^3 + \frac{(u-\ell)^3}{36n^2}$.

Proof.

$$\mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi})) = \frac{1}{12}\left[\sum_{i=1}^{n} \xi_i \xi_{i-1}(\xi_{i-1} - \xi_i) + u^3 - 2u^2\ell + 2u\ell^2 - \ell^3\right]$$

# Analysis of quadratic function

**Theorem**

Given $n \geq 2$, $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, we have that $\xi_i := \ell + \frac{i}{n}(u-\ell)$, for $i = 1, \ldots, n-1$, is the unique minimizer of $\mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$, and the minimum volume is $\frac{1}{18}(u-\ell)^3 + \frac{(u-\ell)^3}{36n^2}$.

Proof.

$$\mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi})) = \frac{1}{12}\left[\sum_{i=1}^{n} \xi_i\xi_{i-1}(\xi_{i-1}-\xi_i) + u^3 - 2u^2\ell + 2u\ell^2 - \ell^3\right]$$

$$\frac{\partial^2 \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))}{\partial \xi_i^2} = \frac{1}{6}(\xi_{i+1} - \xi_{i-1}), \text{ for } i = 1, \ldots, n-1,$$

$$\frac{\partial^2 \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}} = \frac{1}{6}(\xi_i - \xi_{i+1}), \text{ for } i = 1, \ldots, n-2.$$

$\nabla^2 \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$ is tridiagonal and diagonally dominant
$\Rightarrow \nabla^2 \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi})) \succeq 0 \;\Rightarrow\; \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$ is convex.

# Analysis of quadratic function

**Theorem**

Given $n \geq 2$, $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, we have that $\xi_i := \ell + \frac{i}{n}(u - \ell)$, for $i = 1, \ldots, n-1$, is the unique minimizer of $\mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))$, and the minimum volume is $\frac{1}{18}(u-\ell)^3 + \frac{(u-\ell)^3}{36n^2}$.

Proof. The global minimizer satisfies $\nabla \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi})) = 0$.

$$\frac{\partial \mathrm{vol}(\hat{U}_2^*(\boldsymbol{\xi}))}{\partial \xi_i} = \frac{1}{12}(\xi_{i+1} - \xi_{i-1})(2\xi_i - \xi_{i+1} - \xi_{i-1}), \text{ for } i = 1, \ldots, n-1$$

Therefore $2\xi_i - \xi_{i+1} - \xi_{i-1} = 0$ for $i = 1, \ldots, n-1$. Solving these equations gives us the equally spaced points.

# Analysis of non-quadratic convex power functions

$$\operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = \frac{1}{6}\sum_{i=1}^{n}\left|\det\begin{pmatrix}\tau_0 & \tau_i & \tau_{i+1}\\ g(\tau_0) & g(\tau_i) & g(\tau_{i+1})\\ 1 & 1 & 1\end{pmatrix}\right|$$

$$= -\frac{(p-1)^2}{6p}\sum_{i=1}^{n}\frac{(\xi_i^p - \xi_{i-1}^p)^2}{\xi_i^{p-1} - \xi_{i-1}^{p-1}} + \frac{1}{6}((p-1)u^{p+1} - u^p\ell + u\ell^p - (p-1)\ell^p)$$

$$= -\frac{(p-1)^2}{6p}\sum_{i=1}^{n}\frac{\xi_{i-1}^{p-1}\xi_i^{p-1}(\xi_i - \xi_{i-1})^2}{\xi_i^{p-1} - \xi_{i-1}^{p-1}}$$

$$+ \frac{(p-1)(u^{p+1} - \ell^{p+1}) - p(u^p\ell - u\ell^p)}{6p}$$

$\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is still tridiagonal

# Analysis of non-quadratic convex power functions

**Theorem**

For $p > 1$, and $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, if $\boldsymbol{\xi}$ satisfies $\nabla \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = 0$, then $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is positive definite.

# Analysis of non-quadratic convex power functions

**Theorem**

For $p > 1$, and $0 \leq \xi_0 := \ell < \xi_1 < \cdots < \xi_{n-1} < u =: \xi_n$, if $\boldsymbol{\xi}$ satisfies $\nabla \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = 0$, then $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is positive definite.

Any stationary point of $\operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is a strict local minimizer.

# Proof of the local minimizer theorem

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}} = -\frac{(p-1)^2}{3p} \frac{\xi_i^{p-2}\xi_{i+1}^{p-2}[(p-1)\xi_{i+1}^p + \xi_i^p - p\xi_{i+1}^{p-1}\xi_i)][\xi_{i+1}^p + (p-1)\xi_i^p - p\xi_{i+1}\xi_i^{p-1}]}{(\xi_{i+1}^{p-1} - \xi_i^{p-1})^3}$$

# Proof of the local minimizer theorem

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}} = -\frac{(p-1)^2}{3p} \frac{\xi_i^{p-2}\xi_{i+1}^{p-2}[(p-1)\xi_{i+1}^p + \xi_i^p - p\xi_{i+1}^{p-1}\xi_i)][\xi_{i+1}^p + (p-1)\xi_i^p - p\xi_{i+1}\xi_i^{p-1}]}{(\xi_{i+1}^{p-1} - \xi_i^{p-1})^3}$$

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i^2} = \frac{p}{\xi_i} \frac{\partial \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i} - \frac{\xi_{i-1}}{\xi_i} \frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i-1}} - \frac{\xi_{i+1}}{\xi_i} \frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}}$$

# Proof of the local minimizer theorem

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}} = -\frac{(p-1)^2}{3p} \frac{\xi_i^{p-2}\xi_{i+1}^{p-2}[(p-1)\xi_{i+1}^p + \xi_i^p - p\xi_{i+1}^{p-1}\xi_i)][\xi_{i+1}^p + (p-1)\xi_i^p - p\xi_{i+1}\xi_i^{p-1}]}{(\xi_{i+1}^{p-1} - \xi_i^{p-1})^3}$$

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i^2} = \frac{p}{\xi_i}\frac{\partial \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i} - \frac{\xi_{i-1}}{\xi_i}\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i-1}} - \frac{\xi_{i+1}}{\xi_i}\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}}$$

- If $\boldsymbol{\xi}$ satisfies $\nabla \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = 0$, then $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is an $(n-1) \times (n-1)$ symmetric tridiagonal matrix with off-diagonal elements $-b_1, \ldots, -b_{n-2}$ and diagonal elements $a_1, \ldots, a_{n-1}$ where $a_i := \frac{\xi_{i-1}}{\xi_i} b_{i-1} + \frac{\xi_{i+1}}{\xi_i} b_i$, $b_0 \geq 0$, $b_i > 0$

# Proof of the local minimizer theorem

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}} = -\frac{(p-1)^2}{3p} \frac{\xi_i^{p-2}\xi_{i+1}^{p-2}[(p-1)\xi_{i+1}^p + \xi_i^p - p\xi_{i+1}^{p-1}\xi_i)][\xi_{i+1}^p + (p-1)\xi_i^p - p\xi_{i+1}\xi_i^{p-1}]}{(\xi_{i+1}^{p-1} - \xi_i^{p-1})^3}$$

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i^2} = \frac{p}{\xi_i} \frac{\partial \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i} - \frac{\xi_{i-1}}{\xi_i} \frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i-1}} - \frac{\xi_{i+1}}{\xi_i} \frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}}$$

- If $\boldsymbol{\xi}$ satisfies $\nabla \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = 0$, then $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is an $(n-1) \times (n-1)$ symmetric tridiagonal matrix with off-diagonal elements $-b_1, \ldots, -b_{n-2}$ and diagonal elements $a_1, \ldots, a_{n-1}$ where $a_i := \frac{\xi_{i-1}}{\xi_i} b_{i-1} + \frac{\xi_{i+1}}{\xi_i} b_i$, $b_0 \geq 0$, $b_i > 0$
- Notice that $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = \lambda e_1 e_1^\top + M$, where $\lambda = \frac{\xi_0 b_0}{\xi_1} \geq 0$, $M := PDP^\top$, $D := \operatorname{diag}(\frac{\xi_2}{\xi_1} b_1, \frac{\xi_3}{\xi_2} b_2, \ldots, \frac{\xi_n}{\xi_{n-1}} b_{n-1})$, and $P = [p_{ij}]$ is a lower-triangular matrix with $p_{ii} = 1$, $p_{i,i-1} = -\frac{\xi_{i-1}}{\xi_i}$ and 0 otherwise.

# Proof of the local minimizer theorem

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}} = -\frac{(p-1)^2}{3p} \frac{\xi_i^{p-2}\xi_{i+1}^{p-2}[(p-1)\xi_{i+1}^p + \xi_i^p - p\xi_{i+1}^{p-1}\xi_i)][\xi_{i+1}^p + (p-1)\xi_i^p - p\xi_{i+1}\xi_i^{p-1}]}{(\xi_{i+1}^{p-1} - \xi_i^{p-1})^3}$$

$$\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i^2} = \frac{p}{\xi_i}\frac{\partial \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i} - \frac{\xi_{i-1}}{\xi_i}\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i-1}} - \frac{\xi_{i+1}}{\xi_i}\frac{\partial^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i \partial \xi_{i+1}}$$

- If $\boldsymbol{\xi}$ satisfies $\nabla \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = 0$, then $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is an $(n-1) \times (n-1)$ symmetric tridiagonal matrix with off-diagonal elements $-b_1, \ldots, -b_{n-2}$ and diagonal elements $a_1, \ldots, a_{n-1}$ where $a_i := \frac{\xi_{i-1}}{\xi_i} b_{i-1} + \frac{\xi_{i+1}}{\xi_i} b_i$, $b_0 \geq 0$, $b_i > 0$
- Notice that $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = \lambda e_1 e_1^\top + M$, where $\lambda = \frac{\xi_0 b_0}{\xi_1} \geq 0$, $M := PDP^\top$, $D := \operatorname{diag}(\frac{\xi_2}{\xi_1} b_1, \frac{\xi_3}{\xi_2} b_2, \ldots, \frac{\xi_n}{\xi_{n-1}} b_{n-1})$, and $P = [p_{ij}]$ is a lower-triangular matrix with $p_{ii} = 1$, $p_{i,i-1} = -\frac{\xi_{i-1}}{\xi_i}$ and 0 otherwise.
- Therefore $\nabla^2 \operatorname{vol}(\hat{U}_p^*(\boldsymbol{\xi})) \succ 0$

# Analysis of convex power functions (1<p<2)

### Theorem

For $1 < p \leq 2$, and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is strictly convex in $(\xi_1, \ldots, \xi_{n-1})$.

# Analysis of convex power functions (1<p<2)

**Theorem**

For $1 < p \leq 2$, and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is strictly convex in $(\xi_1, \ldots, \xi_{n-1})$.

**Corollary**

For $1 < p \leq 2$ and fixed $\ell,\ u,\ n$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ has a unique minimizer satisfying $\ell < \xi_1 < \cdots < \xi_{n-1} < u$.

# Analysis of convex power functions (1<p<2)

**Theorem**

For $1 < p \leq 2$, and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is strictly convex in $(\xi_1, \ldots, \xi_{n-1})$.

**Corollary**

For $1 < p \leq 2$ and fixed $\ell,\ u,\ n$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ has a unique minimizer satisfying $\ell < \xi_1 < \cdots < \xi_{n-1} < u$.

- $\nabla^2 \mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is not guaranteed to be diagonally dominant. For example, $p = 1.5$, $n = 2$, $\boldsymbol{\xi} = (0, 0.2, 0.8, 1)$, $\nabla^2 \mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi})) \approx \begin{bmatrix} 0.1366 & -0.0621 \\ -0.0621 & 0.0587 \end{bmatrix}$

# Analysis of convex power functions (1<p<2)

**Theorem**

For $1 < p \leq 2$, and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is strictly convex in $(\xi_1, \ldots, \xi_{n-1})$.

**Corollary**

For $1 < p \leq 2$ and fixed $\ell,\ u,\ n$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ has a unique minimizer satisfying $\ell < \xi_1 < \cdots < \xi_{n-1} < u$.

- $\nabla^2 \mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is not guaranteed to be diagonally dominant. For example, $p = 1.5$, $n = 2$, $\boldsymbol{\xi} = (0, 0.2, 0.8, 1)$, $\nabla^2 \mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi})) \approx \begin{bmatrix} 0.1366 & -0.0621 \\ -0.0621 & 0.0587 \end{bmatrix}$
- We will apply a result from Anđelić and Da Fonseca [2011]: if $a_i > 0$ and $\left\{ \frac{b_i^2}{a_i a_{i+1}} \right\}_{i=1}^{n-2}$ is a *chain sequence*, then the tridiagonal matrix is positive definite

# Analysis of convex power functions (p>2)

- If $p > 2$, $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is neither convex nor quasi-convex in general.

# Analysis of convex power functions (p>2)

- If $p > 2$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is neither convex nor quasi-convex in general.

- $\frac{\partial \, \mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i} =$
$$-\frac{(p-1)\xi_i^{p-2}}{6p}\left(\left(\frac{\xi_i^p+(p-1)\xi_{i+1}^p-p\xi_i\xi_{i+1}^{p-1}}{\xi_{i+1}^{p-1}-\xi_i^{p-1}}\right)^2-\left(\frac{\xi_i^p+(p-1)\xi_{i-1}^p-p\xi_i\xi_{i-1}^{p-1}}{\xi_{i-1}^{p-1}-\xi_i^{p-1}}\right)^2\right)$$

# Analysis of convex power functions (p>2)

- If $p > 2$, $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is neither convex nor quasi-convex in general.
- $\frac{\partial \, \text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))}{\partial \xi_i} =$
  $-\frac{(p-1)\xi_i^{p-2}}{6p} \left( \left( \frac{\xi_i^p + (p-1)\xi_{i+1}^p - p\xi_i\xi_{i+1}^{p-1}}{\xi_{i+1}^{p-1} - \xi_i^{p-1}} \right)^2 - \left( \frac{\xi_i^p + (p-1)\xi_{i-1}^p - p\xi_i\xi_{i-1}^{p-1}}{\xi_{i-1}^{p-1} - \xi_i^{p-1}} \right)^2 \right)$
- $\nabla \, \text{vol}(\hat{U}_p^*(\boldsymbol{\xi})) = 0$ is equivalent to $F(\boldsymbol{\xi}) = 0$, where $F(\boldsymbol{\xi}) = [F_1(\boldsymbol{\xi}), F_2(\boldsymbol{\xi}), \ldots, F_{n-1}(\boldsymbol{\xi})]^\top$,

$$F_i(\boldsymbol{\xi}) := -\frac{\xi_i^p + (p-1)\xi_{i+1}^p - p\xi_i\xi_{i+1}^{p-1}}{\xi_{i+1}^{p-1} - \xi_i^{p-1}} + \frac{\xi_i^p + (p-1)\xi_{i-1}^p - p\xi_i\xi_{i-1}^{p-1}}{\xi_i^{p-1} - \xi_{i-1}^{p-1}}.$$

# Analysis of convex power functions (p>2)

**Lemma**

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is nonnegative and $F_i(\boldsymbol{\xi})$ is concave.

# Analysis of convex power functions (p>2)

**Lemma**

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is nonnegative and $F_i(\boldsymbol{\xi})$ is concave.

**Theorem**

If $p > 2$, there exists a unique $\boldsymbol{\xi}^*$ $(\ell < \xi_1^* < \cdots < \xi_{n-1}^* < u)$ such that $F(\boldsymbol{\xi}^*) = 0$.

# Analysis of convex power functions (p>2)

### Lemma

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is nonnegative and $F_i(\boldsymbol{\xi})$ is concave.

### Theorem

If $p > 2$, there exists a unique $\boldsymbol{\xi}^*$ ($\ell < \xi_1^* < \cdots < \xi_{n-1}^* < u$) such that $F(\boldsymbol{\xi}^*) = 0$.

### Corollary

For $p > 2$ and fixed $\ell,\ u,\ n$, $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ has a unique minimizer satisfying $\ell < \xi_1 < \cdots < \xi_{n-1} < u$.

# Proof of the uniqueness theorem

**Lemma**

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is non-negative and $F_i(\boldsymbol{\xi})$ is concave.

- Suppose that $F(\boldsymbol{\xi}^1) = F(\boldsymbol{\xi}^2) = 0$.

# Proof of the uniqueness theorem

**Lemma**

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is non-negative and $F_i(\boldsymbol{\xi})$ is concave.

- Suppose that $F(\boldsymbol{\xi}^1) = F(\boldsymbol{\xi}^2) = 0$.
- Because $F_i(\boldsymbol{\xi})$ is concave,

$$0 = F(\boldsymbol{\xi}^1) - F(\boldsymbol{\xi}^2) \leq F'(\boldsymbol{\xi}^2)(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2)$$

# Proof of the uniqueness theorem

**Lemma**

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is non-negative and $F_i(\boldsymbol{\xi})$ is concave.

- Suppose that $F(\boldsymbol{\xi}^1) = F(\boldsymbol{\xi}^2) = 0$.
- Because $F_i(\boldsymbol{\xi})$ is concave,

$$0 = F(\boldsymbol{\xi}^1) - F(\boldsymbol{\xi}^2) \leq F'(\boldsymbol{\xi}^2)(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2)$$

- Because $[F'(\boldsymbol{\xi}^1)]^{-1}$ and $[F'(\boldsymbol{\xi}^2)]^{-1}$ are nonnegative,

$$\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2 = [F'(\boldsymbol{\xi}^2)]^{-1}(F'(\boldsymbol{\xi}^2)(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2)) \geq 0$$

# Proof of the uniqueness theorem

## Lemma

If $p > 2$ and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, then $[F'(\boldsymbol{\xi})]^{-1}$ is non-negative and $F_i(\boldsymbol{\xi})$ is concave.

- Suppose that $F(\boldsymbol{\xi}^1) = F(\boldsymbol{\xi}^2) = 0$.
- Because $F_i(\boldsymbol{\xi})$ is concave,

$$0 = F(\boldsymbol{\xi}^1) - F(\boldsymbol{\xi}^2) \leq F'(\boldsymbol{\xi}^2)(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2)$$

- Because $[F'(\boldsymbol{\xi}^1)]^{-1}$ and $[F'(\boldsymbol{\xi}^2)]^{-1}$ are nonnegative,

$$\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2 = [F'(\boldsymbol{\xi}^2)]^{-1}(F'(\boldsymbol{\xi}^2)(\boldsymbol{\xi}^1 - \boldsymbol{\xi}^2)) \geq 0$$

- Similarly $\boldsymbol{\xi}^2 - \boldsymbol{\xi}^1 \geq 0$. Thus $\boldsymbol{\xi}^1 = \boldsymbol{\xi}^2$.

# Monotonicity with respect to p

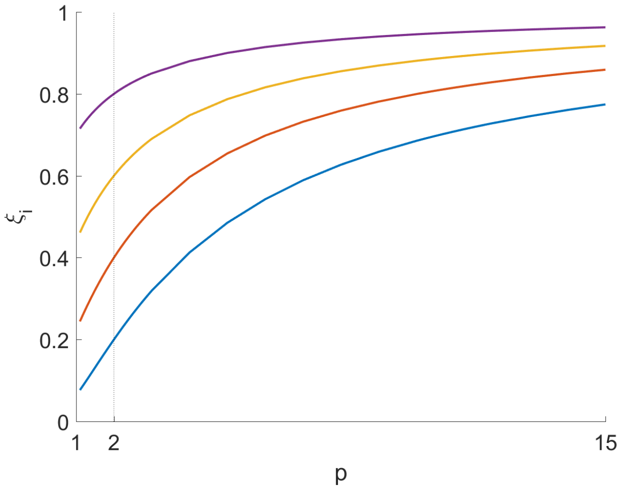



Figure: minimizing $\boldsymbol{\xi}$ for varying $p$ ($n = 5$, $\ell = 0$, $u = 1$).

# Monotonicity with respect to p

### Theorem

For fixed $\ell$ and $u$, and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, suppose that $\boldsymbol{\xi} = (\ell, \xi_1, \ldots, \xi_{n-1}, u)$ minimizes $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$. Then $\xi_i$ $(i = 1, 2, \ldots, n-1)$ is increasing in $p$ on $(1, \infty)$.

# Monotonicity with respect to p

**Theorem**

For fixed $\ell$ and $u$, and $\ell < \xi_1 < \cdots < \xi_{n-1} < u$, suppose that $\boldsymbol{\xi} = (\ell, \xi_1, \ldots, \xi_{n-1}, u)$ minimizes $\mathrm{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$. Then $\xi_i$ $(i = 1, 2, \ldots, n-1)$ is increasing in $p$ on $(1, \infty)$.

Proof Idea.

- Use Implicit function theorem
- Use the following inequality: for $x \in (0,1) \cup (1, \infty)$,

$$\phi(x) := p(p-1)(1-x)x^{p-1} \log x + (x^{p-1} - 1)(x^p - 1) > 0$$

# Conclusion

- We give a way to calulate the volume of the perspective relaxation of piecewise-linear under-estimators $\text{vol}(\hat{U}_f^*(\boldsymbol{\xi}))$
- We prove that the minimizer of $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is unique for convex power functions given the number of linearization points
- We provide some properties of $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ and the minimizers to help choose the linearization points
- We establish that the Newton's method to find stationary point of $\text{vol}(\hat{U}_p^*(\boldsymbol{\xi}))$ is monotonic

# Thank You For Your Attention

Jon Lee and Walter D. Morris, Jr. Geometric comparison of combinatorial polytopes. *Discrete Applied Mathematics*, 55(2):163–182, 1994.

Jon Lee, Daphne Skipper, and Emily Speakman. Gaining or losing perspective. `https://arxiv.org/abs/2001.01435` (journal version of Lee et al. [2020]), 2019.

Jon Lee, Daphne Skipper, and Emily Speakman. Gaining or losing perspective. In Hoai An Le Thi, Hoai Minh Le, and Tao Pham Dinh, editors, *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, pages 387–397. Springer, 2020.

Milica Anđelić and CM Da Fonseca. Sufficient conditions for positive definiteness of tridiagonal matrices revisited. *Positivity*, 15(1):155–159, 2011.

# Chain Sequence

- There exists a parameter sequence $\{c_i\}_{i=0}^{n-2}$ such that $\frac{b_i^2}{a_i a_{i+1}} = c_i(1 - c_{i-1})$ with $0 \leq c_0 < 1$ and $0 < c_i < 1$ for $i \geq 1$
- If $\{\alpha_i\}$ is a chain sequence, and $0 < \beta_i \leq \alpha_i$, then $\{\beta_i\}$ is also a chain sequence.