Convex Optimization and Applications to Image Registration

Brian Knight

Advisor: Dana Paquin

Cal Poly

Spring 2020

We've all got problems

In everyday life, many of the problems we seek to solve are, at heart, optimization problems. *We seek to maximize or minimize some quantity* in the presence of constraints.

$$\begin{array}{ll} \text{minimize} & f(X) \\ \text{subject to} & \text{some restrictions on } X. \end{array}$$

We've all got problems

In everyday life, many of the problems we seek to solve are, at heart, optimization problems. *We seek to maximize or minimize some quantity* in the presence of constraints.

$$\begin{array}{ll} \text{minimize} & f(X) \\ \text{subject to} & \text{some restrictions on } X. \end{array}$$

For the most part, these problems are *easy to state,* and *hard to solve.*

For example, most people are attempting to solve some sort of optimization problem related to happiness or joy or satisfaction in life. We want to maximize this quantity that is subject to the many different constraints faced day to day.

$$\begin{aligned} &\text{maximize} \quad \text{Joy} \\ &\text{subject to} \quad \text{Responsibilities.} \end{aligned}$$

As mathematicians, though, we seek to conquer vagueness. We want to know exactly what it means to be an optimization problem.

As mathematicians, though, we seek to conquer vagueness. We want to know exactly what it means to be an optimization problem.

Definition (*Mathematical Optimization Problem*)

A *mathematical optimization problem* is a problem of the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f_0(x) \quad (\textit{objective function}) \\ \text{subject to} \quad & f_i(x) \leq b_i \text{ for } i = 1, 2, \ldots, m, \\ & h_i(x) = c_i \text{ for } i = 1, 2, \ldots, p, \\ & \qquad\qquad (\textit{constraints}) \end{aligned}$$

where $b_i, \in \mathbb{R}$ for $i = 1, 2, \ldots, m$, $c_i, \in \mathbb{R}$ for $i = 1, 2, \ldots, p$.

Familiar Optimization Examples

Example (Minimizing a Quadratic)

$$\text{minimize}_{x \in \mathbb{R}^n} \, f(x) = x^T P x + q^T x + c$$

Familiar Optimization Examples

Example (Minimizing a Quadratic)

$$\text{minimize}_{x\in\mathbb{R}^n}\ f(x) = x^T P x + q^T x + c$$

Example (Constrained Least Squares)

$$\begin{aligned}&\text{minimize}_{x\in\mathbb{R}^n}\ \|Ax - b\| \\ &\text{subject to } b_{lb} \preccurlyeq x \preccurlyeq b_{ub}\end{aligned}$$

Familiar Optimization Examples

Example (Minimizing a Quadratic)

$$\text{minimize}_{x \in \mathbb{R}^n} \; f(x) = x^T P x + q^T x + c$$

Example (Constrained Least Squares)

$$\begin{aligned} &\text{minimize}_{x \in \mathbb{R}^n} \; \|Ax - b\| \\ &\text{subject to } b_{lb} \preccurlyeq x \preccurlyeq b_{ub} \end{aligned}$$

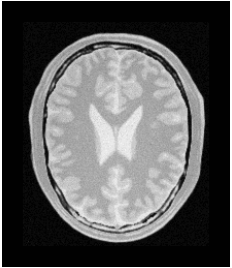
Example (Portfolio Optimization)

$$\begin{aligned} \text{maximize}_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{subject to} \quad & x \succcurlyeq 0 \\ & \mathbf{1}^T x = 1 \end{aligned}$$

The Image Registration Problem

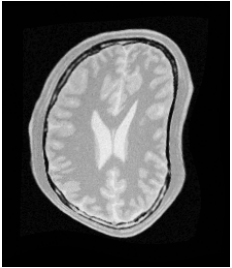
We will be particularly interested in the *image registration problem*, which is formulated mathematically as follows

> *Given two images A and B, a distance function D, and a transformation space τ, find the transformation Φ such that*
>
> $$\Phi = \mathrm{argmin}_{\varphi \in \tau} D(A, \varphi(B)).$$



Base, B

$$\xrightarrow{\varphi(B)}$$



Target, A

This problem has been widely studied and there are many different approaches to determining an optimal trasnformation. It has applications in

This problem has been widely studied and there are many different approaches to determining an optimal trasnformation. It has applications in

- Medical Imaging

This problem has been widely studied and there are many different approaches to determining an optimal trasnformation. It has applications in

- Medical Imaging
- Computer Vision

The Image Registration Problem

This problem has been widely studied and there are many different approaches to determining an optimal trasnformation. It has applications in

- Medical Imaging
- Computer Vision
- Autonomous Vehicle Navigation

The Image Registration Problem

This problem has been widely studied and there are many different approaches to determining an optimal trasnformation. It has applications in

- Medical Imaging
- Computer Vision
- Autonomous Vehicle Navigation
- Data Analysis of Images (*e.g.* satellite data)

The Image Registration Problem

This problem has been widely studied and there are many different approaches to determining an optimal trasnformation. It has applications in

- Medical Imaging
- Computer Vision
- Autonomous Vehicle Navigation
- Data Analysis of Images (*e.g.* satellite data)

Our goal is to state the image registration problem as an optimization problem with a unique solution.

The image registration problem can be easily cast as a standard optimization problem in the following way:

$$\underset{\varphi \in \tau}{\text{minimize}} \quad D(A, \varphi(B)).$$

The Image Registration Problem

The image registration problem can be easily cast as a standard optimization problem in the following way:

$$\underset{\varphi \in \tau}{\text{minimize}} \quad D(A, \varphi(B)).$$

- $A, B : \mathbb{R}^{m \times n} \to \mathbb{R}$,
  $A(x, y) =$ an intensity value in $\{0, \ldots, 255\}$ (a standard 8-bit image).
  $A_{xy} := A(x, y)$.

The Image Registration Problem

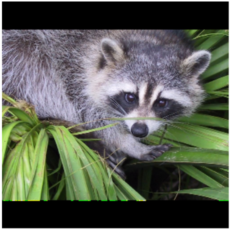
The image registration problem can be easily cast as a standard optimization problem in the following way:

$$\underset{\varphi \in \tau}{\text{minimize}} \quad D(A, \varphi(B)).$$

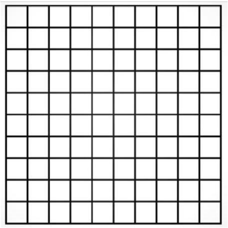
- $A, B : \mathbb{R}^{m \times n} \to \mathbb{R}$,
  $A(x, y) =$ an intensity value in $\{0, \ldots, 255\}$ (a standard 8-bit image).
  $A_{xy} := A(x, y)$.
- Color images (RGB) will be functions into $\mathbb{R}^3$ with intensity values as triples in $\{0, \ldots, 255\}^3$.

The Image Registration Problem

For this talk we consider each $\varphi \in \tau$ to be a spacial transformation on coordinates in the physical space of the image:



$\xrightarrow{\varphi}$



$\xrightarrow{\varphi}$

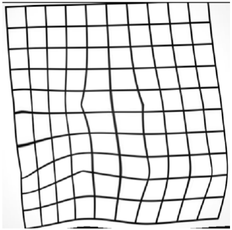


Image Grid

Transformed Image Grid

It is important to notice that our minimization problem

$$\underset{\varphi \in \tau}{\text{minimize}} \quad D(A, \varphi(B)),$$

is highly dependent upon this distance function D.

A standard choice for this function is the *Mean Squares Metric*.

It is important to notice that our minimization problem

$$\underset{\varphi \in \tau}{\text{minimize}} \quad D(A, \varphi(B)),$$

is highly dependent upon this distance function D.

A standard choice for this function is the *Mean Squares Metric*.

Definition (*Mean Squares Metric*)

Given two images A and B, the *Mean Squares Metric* is the function

$$MS(A, B) = \frac{1}{mn} \sum_{x=1}^{n} \sum_{y=1}^{m} (A_{xy} - B_{xy})^2,$$

where A_{xy} and B_{xy} represent the intensity values at pixel (x, y) of the images A and B respectively.

Other standard image metrics:

Other standard image metrics:

- Normalized Correlation Metric:

$$NC(A,B) = -\frac{\sum_{i=1}^{mn}(A_i \cdot B_i)}{\sqrt{\sum_{i=1}^{mn} A_i^2 \cdot \sum_{i=1}^{mn} B_i^2}}$$

Other standard image metrics:

- Normalized Correlation Metric:

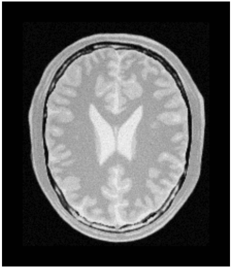
$$NC(A,B) = -\frac{\sum_{i=1}^{mn}(A_i \cdot B_i)}{\sqrt{\sum_{i=1}^{mn} A_i^2 \cdot \sum_{i=1}^{mn} B_i^2}}$$

- Mutual Information Metric (allows for multi-modal registration):

$$MI(A,B) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right),$$

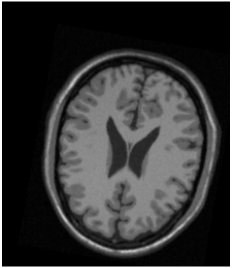
  where $p(x,y)$ is the joint probability of intensities between images, and $p(x), p(y)$ are the probabilities of intensities in image A and B respectively.

Consider a more challenging registration problem of registering two medical images of different modality (*e.g.* CT vs. MRI):



Base, B

$$\xrightarrow{\varphi(B)}$$



Target, A

It is clear that the Mean Squares Metric (MS) will find that, for any two images A and B,

$$MS(A, B) = 0 \Leftrightarrow A = B.$$

The Image Registration Problem

It is clear that the Mean Squares Metric (MS) will find that, for any two images A and B,

$$MS(A, B) = 0 \Leftrightarrow A = B.$$

- What if $A \neq B$? If there is some third image C, and

$$MS(A, B) \leq MS(A, C),$$

  does that mean that B is ‘closer’ to A than C is?

The Image Registration Problem

It is clear that the Mean Squares Metric (MS) will find that, for any two images A and B,

$$MS(A, B) = 0 \Leftrightarrow A = B.$$

- What if $A \neq B$? If there is some third image C, and

  $$MS(A, B) \leq MS(A, C),$$

  does that mean that B is 'closer' to A than C is?
- If $\varphi, \sigma \in \tau$ are two transformations, and

  $$MS(A, \varphi(B)) \leq MS(A, \sigma(B)),$$

  is φ a 'better' transformation than σ?

More uncertainties:

More uncertainties:

- What if we cannot find a transformation φ that maps the base B to A exactly? So $D(A, \varphi(B))$ never reaches zero.

More uncertainties:

- What if we cannot find a transformation φ that maps the base B to A exactly? So $D(A, \varphi(B))$ never reaches zero.
- If we find a candidate for φ, how do we know it is the best possible solution?

More uncertainties:

- What if we cannot find a transformation φ that maps the base B to A exactly? So $D(A, \varphi(B))$ never reaches zero.
- If we find a candidate for φ, how do we know it is the best possible solution?

The issues that arise in solving the Image Registration problem are not unique to it. In all optimization problems we are constantly running into the issue that

$$\text{local minima} \;\not\Leftrightarrow\; \text{global minima} ☹.$$

Again the mathematician is faced with an ambiguity that must be reconciled.

The Promised Land (Reconciliation of the Minima)

Again the mathematician is faced with an ambiguity that must be reconciled.

- Is there a land out there in which all problems have globally optimal solutions?

The Promised Land (Reconciliation of the Minima)

Again the mathematician is faced with an ambiguity that must be reconciled.

- Is there a land out there in which all problems have globally optimal solutions?
- If so, what sort of structure does this land have?

Again the mathematician is faced with an ambiguity that must be reconciled.

- Is there a land out there in which all problems have globally optimal solutions?
- If so, what sort of structure does this land have?

With a sufficiently nice set of conditions on our optimization problems we can resolve the local minima issue and ensure that if a local minimum exists it is a global minimum.

$$\text{local minima} \;\Leftrightarrow\; \text{global minima} \; ☺.$$

The Promised Land (Reconciliation of the Minima)

Again the mathematician is faced with an ambiguity that must be reconciled.

- Is there a land out there in which all problems have globally optimal solutions?
- If so, what sort of structure does this land have?

With a sufficiently nice set of conditions on our optimization problems we can resolve the local minima issue and ensure that if a local minimum exists it is a global minimum.

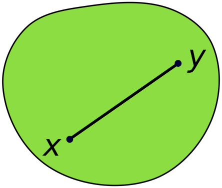
$$\text{local minima} \;\Leftrightarrow\; \text{global minima} \;☺.$$

The central idea in all of these conditions is that of *convexity*.

Definition (*convex set*)

A set $C \subseteq \mathbb{R}^n$ is *convex* if for any two points $x_1, x_2 \in C$, the line segment between them lies in C. i.e.

$$\forall\, \theta \in [0,1], \quad \theta x_1 + (1-\theta) x_2 \in C$$



Cartoon of a Convex Set in $\mathbb{R}^2$

Example

Give any norm $\|\cdot\|$ on $\mathbb{R}^n$, the closed norm ball centered at $x_c \in \mathbb{R}^n$ of radius $r > 0$ defined by

$$B(x_c, r) = \{y \in \mathbb{R}^n \mid \|x_c - y\| \leq r\},$$

is a convex set.

Example

Give any norm $\|\cdot\|$ on $\mathbb{R}^n$, the closed norm ball centered at $x_c \in \mathbb{R}^n$ of radius $r > 0$ defined by

$$B(x_c, r) = \{y \in \mathbb{R}^n \mid \|x_c - y\| \leq r\},$$

is a convex set.

Example

The hyperplane parallel to $a \in \mathbb{R}^n$ at b is the set

$$\{x \in \mathbb{R}^n \mid a^T x = b\}.$$

A hyperplane is a convex set, and each hyperplane splits $\mathbb{R}^n$ into two *halfspaces*

$$\{x \in \mathbb{R}^n \mid a^T x \leq b\} \text{ and } \{x \in \mathbb{R}^n \mid a^T x \geq b\},$$

both of which are convex sets.

Some nice properties:

Some nice properties:

- A set $C \subseteq \mathbb{R}^n$ is convex if and only if its intersection with any line $L \in \mathbb{R}^n$ is convex.

Some nice properties:

- A set $C \subseteq \mathbb{R}^n$ is convex if and only if its intersection with any line $L \in \mathbb{R}^n$ is convex.
- Let $\{C_\lambda\}_{\lambda \in \Lambda}$ be an arbitrary collection of convex sets. Then

$$\bigcap_{\lambda \in \Lambda} C_\lambda \text{ is convex.}$$

Some nice properties:

- A set $C \subseteq \mathbb{R}^n$ is convex if and only if its intersection with any line $L \in \mathbb{R}^n$ is convex.
- Let $\{C_\lambda\}_{\lambda \in \Lambda}$ be an arbitrary collection of convex sets. Then

$$\bigcap_{\lambda \in \Lambda} C_\lambda \text{ is convex.}$$

Proof: Let x, y to be two points in the intersection.

Some nice properties:

- A set $C \subseteq \mathbb{R}^n$ is convex if and only if its intersection with any line $L \in \mathbb{R}^n$ is convex.
- Let $\{C_\lambda\}_{\lambda \in \Lambda}$ be an arbitrary collection of convex sets. Then

$$\bigcap_{\lambda \in \Lambda} C_\lambda \text{ is convex.}$$

Proof: Let x, y to be two points in the intersection. Then $x, y \in C_\lambda$ for all $\lambda \in \Lambda$,

Some nice properties:

- A set $C \subseteq \mathbb{R}^n$ is convex if and only if its intersection with any line $L \in \mathbb{R}^n$ is convex.
- Let $\{C_\lambda\}_{\lambda \in \Lambda}$ be an arbitrary collection of convex sets. Then

$$\bigcap_{\lambda \in \Lambda} C_\lambda \text{ is convex.}$$

Proof: Let x, y to be two points in the intersection. Then $x, y \in C_\lambda$ for all $\lambda \in \Lambda$, so

$$\theta x + (1-\theta) y \in C_\lambda$$

for any $\theta \in [0,1]$ and for each $\lambda \in \Lambda$.

Some nice properties:

- A set $C \subseteq \mathbb{R}^n$ is convex if and only if its intersection with any line $L \in \mathbb{R}^n$ is convex.
- Let $\{C_\lambda\}_{\lambda \in \Lambda}$ be an arbitrary collection of convex sets. Then

$$\bigcap_{\lambda \in \Lambda} C_\lambda \text{ is convex.}$$

Proof: Let x, y to be two points in the intersection. Then $x, y \in C_\lambda$ for all $\lambda \in \Lambda$, so

$$\theta x + (1-\theta)y \in C_\lambda$$

for any $\theta \in [0,1]$ and for each $\lambda \in \Lambda$. Thus

$$\theta x + (1-\theta)y \in \bigcap_{\lambda \in \Lambda} C_\lambda. \quad \square$$

The result of arbitrary intersections provides a way to construct new convex sets.

- A *polyhedron* in $\mathbb{R}^n$ can be expressed as the the intersection of finitely many hyperplanes and half-spaces, and thus is convex.

The result of arbitrary intersections provides a way to construct new convex sets.

- A *polyhedron* in $\mathbb{R}^n$ can be expressed as the the intersection of finitely many hyperplanes and half-spaces, and thus is convex.
- If we denote $\mathbf{S}^n$ as the set of symmetric matrices in $\mathbb{R}^{n\times n}$, i.e.

$$\mathbf{S}^n = \{X \in \mathbb{R}^{n\times n} | X = X^T\},$$

then the set of positive semidefinite matrices

$$\mathbf{S}^n_+ = \{X \in \mathbf{S}^n | X \succcurlyeq 0\} = \{X \in \mathbb{R}^{n\times n} | x^T X x \geq 0 \, \forall x \in \mathbb{R}^n\},$$

is the intersection of halfspaces in $\mathbf{S}^n$,

$$\bigcap_{x\in\mathbb{R}^n} \left\{X \in \mathbf{S}^n \mid x^T X x \geq 0\right\},$$

and thus is convex.

Another key idea is that of the *convex hull*.

Another key idea is that of the *convex hull*.

Definition (*convex hull*)

Given a set C, the *convex hull* of C, denoted as $\mathbf{conv}\, C$ is the set of all convex combinations of points in C.

$$\mathbf{conv}\, C = \left\{ \sum_{i=0}^{k} \theta_i x_i \,\middle|\, x_i \in C, \theta_i \geq 0, i = 1, 2, \ldots, k, \sum_{i}^{k} \theta_i = 1 \right\}.$$

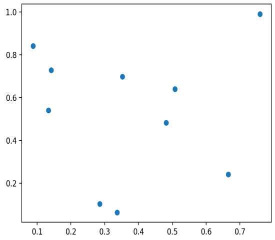
Basics of Convexity

Another key idea is that of the *convex hull*.

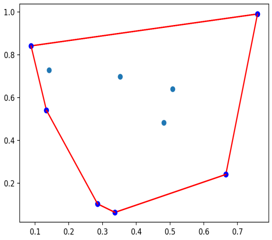
Definition (*convex hull*)

Given a set C, the *convex hull* of C, denoted as $\mathbf{conv}\, C$ is the set of all convex combinations of points in C.

$$\mathbf{conv}\, C = \left\{ \sum_{i=0}^{k} \theta_i x_i \;\middle|\; x_i \in C, \theta_i \geq 0, i = 1, 2, \ldots, k, \sum_{i}^{k} \theta_i = 1 \right\}.$$



C



conv C

By considering the *epigraph* of a function $f : \mathbb{R} \to \mathbb{R}$, defined as

$$\textbf{epi}\, f = \{(x, t) \in \mathbb{R}^2 \mid f(x) \leq t\}$$

we can immediately relate our definition of a convex set to the codomain of a function.

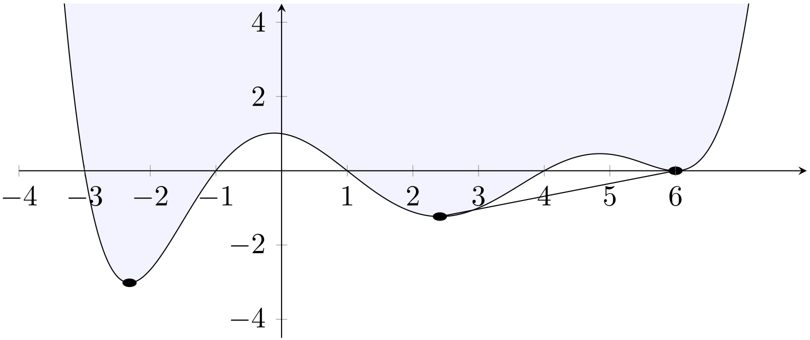


Figure: A non-convex epigraph, local minima $\not\Leftrightarrow$ global minimum

By considering the *epigraph* of a function $f : \mathbb{R} \to \mathbb{R}$, defined as

$$\mathbf{epi}\, f = \{(x, t) \in \mathbb{R}^2 \mid f(x) \leq t\}$$

we can immediately relate our definition of a convex set to the codomain of a function.

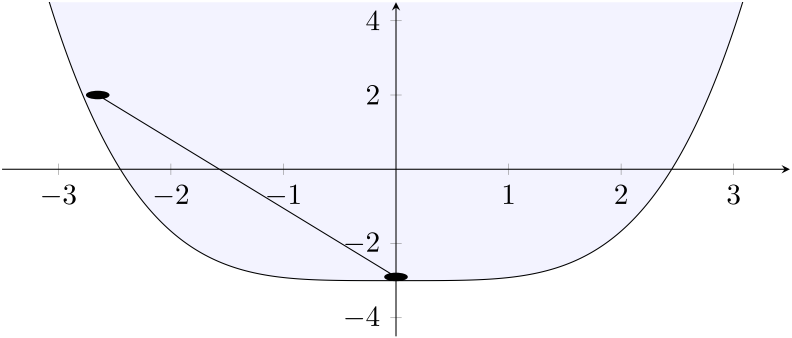


Figure: A convex epigraph, local minima $\Leftrightarrow$ global minimum

From here we have a natural definition of a *convex function*.

Definition (*Convex function*)

A function $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ is *convex* if $C = \mathbf{dom}\, f$ is a convex set and if for all $x, y \in \mathbf{dom}\, f$ and $\theta \in [0, 1]$, the inequality

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

holds.

Basics of Convexity

From here we have a natural definition of a *convex function*.

Definition (*Convex function*)

A function $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ is *convex* if $C = \mathbf{dom}\, f$ is a convex set and if for all $x, y \in \mathbf{dom}\, f$ and $\theta \in [0, 1]$, the inequality

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

holds.

Examples

- The quadratic function $f(x) = ax^2 + bx + c$ is convex $\Leftrightarrow a \geq 0$.

From here we have a natural definition of a *convex function*.

Definition (*Convex function*)

A function $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ is *convex* if $C = \mathbf{dom}\, f$ is a convex set and if for all $x, y \in \mathbf{dom}\, f$ and $\theta \in [0, 1]$, the inequality

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

holds.

Examples

- The quadratic function $f(x) = ax^2 + bx + c$ is convex $\Leftrightarrow a \geq 0$.
- Any affine function $f : \mathbb{R}^n \to \mathbb{R}$ with $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, defined by $f(x) = a^T x + b$ is convex.

From here we have a natural definition of a *convex function*.

Definition (*Convex function*)

A function $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ is *convex* if $C = \textbf{dom}\, f$ is a convex set and if for all $x, y \in \textbf{dom}\, f$ and $\theta \in [0, 1]$, the inequality

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

holds.

Examples

- The quadratic function $f(x) = ax^2 + bx + c$ is convex $\Leftrightarrow a \geq 0$.
- Any affine function $f : \mathbb{R}^n \to \mathbb{R}$ with $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, defined by $f(x) = a^T x + b$ is convex.
- Any norm $\|\cdot\|$ on $\mathbb{R}^n$ is a convex function.

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The α-*sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The α-*sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.
  Proof:

  $f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y)$

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The α-*sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.
  Proof:

  $$f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y) \leq \theta\alpha+(1-\theta)\alpha = \alpha. \quad \square$$

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The α-*sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.
  Proof:

  $$f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y) \leq \theta\alpha+(1-\theta)\alpha = \alpha. \quad \square$$

- The epigraph of f : $\mathbf{epi}\, f = \{(x,t) \in \mathbb{R}^{n+1} \mid f(x) \leq t\}$.

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The α-*sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.
  Proof:

  $$f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y) \leq \theta\alpha+(1-\theta)\alpha = \alpha. \quad \square$$

- The epigraph of f : $\textbf{epi}\, f = \{(x,t) \in \mathbb{R}^{n+1} \mid f(x) \leq t\}$.
  Proof:

  $$(x_1,t_1),(x_2,t_2) \in \textbf{epi}\, f \implies f(x_1) \leq t_1 \text{ and } f(x_2) \leq t_2.$$

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The *α-sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.
  Proof:

  $$f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y) \leq \theta\alpha+(1-\theta)\alpha = \alpha. \quad \square$$

- The epigraph of f : $\mathbf{epi}\, f = \{(x,t) \in \mathbb{R}^{n+1} \mid f(x) \leq t\}$.
  Proof:

  $$(x_1,t_1),(x_2,t_2) \in \mathbf{epi}\, f \implies f(x_1) \leq t_1 \text{ and } f(x_2) \leq t_2.$$

  So, for $\theta \in [0,1]$, $f(\theta x_1 + (1-\theta)x_2) \leq \theta t_1 + (1-\theta)t_2$.

Let $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function. The following sets associated with f are all convex:

- The *α-sublevel set* of f defined as $C_\alpha = \{x \mid f(x) \leq \alpha\}$.
  Proof:
  $$f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y) \leq \theta\alpha+(1-\theta)\alpha = \alpha. \quad \square$$

- The epigraph of f : $\mathbf{epi}\, f = \{(x,t) \in \mathbb{R}^{n+1} \mid f(x) \leq t\}$.
  Proof:
  $$(x_1,t_1),(x_2,t_2) \in \mathbf{epi}\, f \implies f(x_1) \leq t_1 \text{ and } f(x_2) \leq t_2.$$
  So, for $\theta \in [0,1]$, $f(\theta x_1 + (1-\theta)x_2) \leq \theta t_1 + (1-\theta)t_2$. Thus
  $\theta(x_1,t_1) + (1-\theta)(x_2,t_2) \in \mathbf{epi}\, f$. $\square$

Example (Non-negative sums)

Given two convex functions f and g, the sum $f + g$ is convex on $\mathbf{dom}\, f \cap \mathbf{dom}\, g$.

Proof: The domain $\mathbf{dom}\,(f+g) = \mathbf{dom}\, f \cap \mathbf{dom}\, g$ is convex. ✓
So for any $x, y \in \mathbf{dom}\,(f+g)$, and any $\theta \in [0,1]$ it follows

Example (Non-negative sums)

Given two convex functions f and g, the sum $f+g$ is convex on $\mathbf{dom}\, f \cap \mathbf{dom}\, g$.

Proof: The domain $\mathbf{dom}\,(f+g) = \mathbf{dom}\, f \cap \mathbf{dom}\, g$ is convex. ✓ So for any $x, y \in \mathbf{dom}\,(f+g)$, and any $\theta \in [0,1]$ it follows

$$\begin{aligned}(f+g)(\theta x + (1-\theta)y) &\leq \theta f(x) + (1-\theta) f(y) + \\ &\quad \theta g(x) + (1-\theta) g(y) \\ &= \theta(f+g)(x) + (1-\theta)(f+g)(y).\end{aligned}$$ □

Example (Non-negative sums)

Given two convex functions f and g, the sum $f+g$ is convex on $\mathbf{dom}\, f \cap \mathbf{dom}\, g$.

Proof: The domain $\mathbf{dom}\,(f+g) = \mathbf{dom}\, f \cap \mathbf{dom}\, g$ is convex. ✓ So for any $x, y \in \mathbf{dom}\,(f+g)$, and any $\theta \in [0,1]$ it follows

$$\begin{aligned}(f+g)(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) + \\ \theta g(x) + (1-\theta) g(y) \\ = \theta(f+g)(x) + (1-\theta)(f+g)(y). \quad \square\end{aligned}$$

This can be extended to arbitrary finite sums, non-negatively weighted sums, so that

$$\mathcal{F} = \{ f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ is convex } \}$$

forms a convex cone.

With the notion of convex sets and convex functions understood, we can now define a *convex optimization problem*.

Convex Optimization Problems

With the notion of convex sets and convex functions understood, we can now define a *convex optimization problem.*

A standard *convex optimization problem* is an optimization problem of the form:

$$\begin{aligned} \underset{x \in \mathcal{D}}{\text{minimize}} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & a_i^T x = b_i, \quad i = 1, \ldots, p, \end{aligned}$$

where f_i are all convex functions, and $a_i \in \mathbb{R}^n$ (equivalently, $Ax = b$ for $A \in \mathbb{R}^{p \times n}$).

Given our standard convex optimization problem, we have some immediate consequences that inform us our definition is pretty good:

- The domain of the problem, $\mathcal{D}$, is a convex set, as

$$\mathcal{D} = \bigcap_{i=0}^{m} \textbf{dom}\, f_i.$$

Convex Optimization Problems

Given our standard convex optimization problem, we have some immediate consequences that inform us our definition is pretty good:

- The domain of the problem, $\mathcal{D}$, is a convex set, as

$$\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i.$$

- The set of all elements in $\mathcal{D}$ that satisfy all constraints, called the *feasible set*,

$$\mathcal{D} \cap \left(\bigcap_{i=1}^{m} \{x \mid f_i(x) \leq 0\} \right) \cap \left(\bigcap_{i=1}^{p} \{x \mid a_i^T x = b_i\} \right)$$

  is convex, as it is the intersection of $\mathcal{D}$, 0-sublevel sets of convex functions, and hyperplanes.

Convex Optimization Problems

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minima $\tilde{x}$ of f_0 in the feasible set is a global minima of f_0 over the feasible set.

Convex Optimization Problems

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minima $\tilde{x}$ of f_0 in the feasible set is a global minima of f_0 over the feasible set.

Proof.

$\tilde{x}$ feasible, local minima of f_0 implies

$$f_0(x) > f_0(\tilde{x}) \, \forall x \text{ s.t } 0 < \|\tilde{x} - x\| \leq \rho.$$

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minima $\tilde{x}$ of f_0 in the feasible set is a global minima of f_0 over the feasible set.

Proof.

$\tilde{x}$ feasible, local minima of f_0 implies

$$f_0(x) > f_0(\tilde{x}) \, \forall x \text{ s.t } 0 < \|\tilde{x} - x\| \leq \rho.$$

Suppose $f_0(\tilde{x}) > f_0(x^*)$.

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minima $\tilde{x}$ of f_0 in the feasible set is a global minima of f_0 over the feasible set.

Proof.

$\tilde{x}$ feasible, local minima of f_0 implies

$$f_0(x) > f_0(\tilde{x}) \, \forall x \text{ s.t } 0 < \|\tilde{x} - x\| \leq \rho.$$

Suppose $f_0(\tilde{x}) > f_0(x^*)$. $\tilde{x}$ and x^* feasible implies

$$\theta \tilde{x} + (1 - \theta) x^* \text{ feasible } \forall \theta \in [0, 1].$$

Convex Optimization Problems

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minima $\tilde{x}$ of f_0 in the feasible set is a global minima of f_0 over the feasible set.

Proof.

$\tilde{x}$ feasible, local minima of f_0 implies

$$f_0(x) > f_0(\tilde{x}) \, \forall x \text{ s.t } 0 < \|\tilde{x} - x\| \leq \rho.$$

Suppose $f_0(\tilde{x}) > f_0(x^*)$. $\tilde{x}$ and x^* feasible implies

$$\theta\tilde{x} + (1-\theta)x^* \text{ feasible } \forall\theta \in [0,1].$$

Since f_0 is convex we have that

$$\begin{aligned} f_0(\theta\tilde{x} + (1-\theta)x^*) &\leq \theta f_0(\tilde{x}) + (1-\theta)f_0(x^*) \\ &\leq f_0(\tilde{x}) \text{ for any } \theta \in [0,1]. \end{aligned}$$

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minimum $\tilde{x}$ of f_0 in the feasible set is a global minimum of f_0 over the feasible set.

So choose θ to be such that $0 < \|\tilde{x} - (\theta\tilde{x} + (1-\theta)x^*)\| \leq \rho$.

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minimum $\tilde{x}$ of f_0 in the feasible set is a global minimum of f_0 over the feasible set.

So choose θ to be such that $0 < \|\tilde{x} - (\theta\tilde{x} + (1-\theta)x^*)\| \leq \rho$. Then

$$f_0(\theta\tilde{x} + (1-\theta)x^*) > f_0(\tilde{x}),$$

which is a contradiction. □

Theorem (Reconciliation of the Minima)

Give a convex optimization problem, any local minimum $\tilde{x}$ of f_0 in the feasible set is a global minimum of f_0 over the feasible set.

So choose θ to be such that $0 < \|\tilde{x} - (\theta\tilde{x} + (1-\theta)x^*)\| \leq \rho$. Then

$$f_0(\theta\tilde{x} + (1-\theta)x^*) > f_0(\tilde{x}),$$

which is a contradiction. $\square$

- We have just shown that convex optimization problems have the structure that we desired our optimization problems to have; now we set out to solve them.

A Simple Case

Let's start by considering an unconstrained convex optimization problem with f sufficiently differentiable on an open domain:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x).$$

A Simple Case

Let's start by considering an unconstrained convex optimization problem with f sufficiently differentiable on an open domain:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x).$$

- Assuming the problem is solvable, there is some x^* such that $\inf_x f(x) = f(x^*) = p^*$.

A Simple Case

Let's start by considering an unconstrained convex optimization problem with f sufficiently differentiable on an open domain:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x).$$

- Assuming the problem is solvable, there is some x^* such that $\inf_x f(x) = f(x^*) = p^*$.
- This must occur at a point such that $\nabla f(x^*) = 0$.

A Simple Case

Let's start by considering an unconstrained convex optimization problem with f sufficiently differentiable on an open domain:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x).$$

- Assuming the problem is solvable, there is some x^* such that $\inf_x f(x) = f(x^*) = p^*$.
- This must occur at a point such that $\nabla f(x^*) = 0$.
- Since local minima are global minima, a solution to the above equation solves the minimization problem.

A Simple Case

Let's start by considering an unconstrained convex optimization problem with f sufficiently differentiable on an open domain:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x).$$

- Assuming the problem is solvable, there is some x^* such that $\inf_x f(x) = f(x^*) = p^*$.
- This must occur at a point such that $\nabla f(x^*) = 0$.
- Since local minima are global minima, a solution to the above equation solves the minimization problem.
- Sometimes this may be done analytically, but most of the time we use numerical approximations given by various descent methods.

A More Complicated Case

Now we consider the convex optimization problem

$$\begin{array}{ll} \underset{x \in \mathbf{dom}\, f}{\text{minimize}} & f_0(x) \\ \text{subject to} & f_i(x) \le b_i, \quad i = 1, \ldots, m, \end{array}$$

or equivalently

$$\begin{array}{ll} \underset{x \in \mathbf{dom}\, f}{\text{minimize}} & f_0(x) \\ \text{subject to} & f_i(x) - b_i \le 0, \quad i = 1, \ldots, m, \end{array}$$

A More Complicated Case

We make the inequality constraints implicit in the domain of a new, equivalent unconstrained optimization problem:

A More Complicated Case

We make the inequality constraints implicit in the domain of a new, equivalent unconstrained optimization problem:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) + \sum_{i=1}^{m} I_{-}\left(b_i - f_i(x)\right),$$

where $I_{-} : \mathbb{R} \to \overline{\mathbb{R}}$ is defined as

$$I_{-}(u) = \begin{cases} \infty & u < 0 \\ 0 & \text{otherwise.} \end{cases}$$

A More Complicated Case

We make the inequality constraints implicit in the domain of a new, equivalent unconstrained optimization problem:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) + \sum_{i=1}^{m} I_{-}\left(b_i - f_i(x)\right),$$

where $I_{-} : \mathbb{R} \to \overline{\mathbb{R}}$ is defined as

$$I_{-}(u) = \begin{cases} \infty & u < 0 \\ 0 & \text{otherwise.} \end{cases}$$

- If $f_i(x) > b_i$ for any i, $f(x) = \infty$, and so if this problem has a solution it must be feasible.

A More Complicated Case

We make the inequality constraints implicit in the domain of a new, equivalent unconstrained optimization problem:

$$\underset{x \in \textbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) + \sum_{i=1}^{m} I_-\left(b_i - f_i(x)\right),$$

where $I_- : \mathbb{R} \to \overline{\mathbb{R}}$ is defined as

$$I_-(u) = \begin{cases} \infty & u < 0 \\ 0 & \text{otherwise.} \end{cases}$$

- If $f_i(x) > b_i$ for any i, $f(x) = \infty$, and so if this problem has a solution it must be feasible.
- $f(x)$ is still convex (sum of two convex functions), but is clearly no longer differentiable, so descent methods won't work.

To solve this problem, we introduce the *logarithmic barrier function* to approximate a solution to the previous problem:

$$\hat{I}_-(u,t) = -\frac{1}{t}\log(-u), \text{ for } t > 0.$$

To solve this problem, we introduce the *logarithmic barrier function* to approximate a solution to the previous problem:

$$\hat{I}_{-}(u,t) = -\frac{1}{t}\log(-u), \text{ for } t > 0.$$

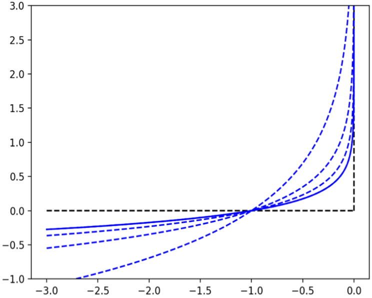


Figure: Plots of the $\hat{I}(u,t)$ for increasing values of t from $t = 1$ to $t = 4$ in blue. Solid blue is $t = 4$. Black dashed line is the function $I_{-}(u)$.

Thus we now have an unconstrained minimization problem with differentiable objective function for some $t > 0$:

Thus we now have an unconstrained minimization problem with differentiable objective function for some $t > 0$:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log\left(-(f_i(x) - b_i)\right),$$

A More Complicated Case

Thus we now have an unconstrained minimization problem with differentiable objective function for some $t > 0$:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log\left(-(f_i(x) - b_i)\right),$$

- $\log(-u)$ is concave, so $-\log(-u)$ is convex. It is also non-decreasing, and the composition of a convex non-decreasing function with a convex function is convex.

A More Complicated Case

Thus we now have an unconstrained minimization problem with differentiable objective function for some $t > 0$:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log\left(-(f_i(x) - b_i)\right),$$

- $\log(-u)$ is concave, so $-\log(-u)$ is convex. It is also non-decreasing, and the composition of a convex non-decreasing function with a convex function is convex.
- As $t \to \infty$, the solution approaches the solution to the original inequality constrained problem.

A More Complicated Case

Thus we now have an unconstrained minimization problem with differentiable objective function for some $t > 0$:

$$\underset{x \in \mathbf{dom}\, f}{\text{minimize}} \quad f(x) = f_0(x) - \frac{1}{t}\sum_{i=1}^{m} \log\left(-(f_i(x) - b_i)\right),$$

- $\log(-u)$ is concave, so $-\log(-u)$ is convex. It is also non-decreasing, and the composition of a convex non-decreasing function with a convex function is convex.
- As $t \to \infty$, the solution approaches the solution to the original inequality constrained problem.
- We solve iteratively in t to ensure a feasible starting position x at as t increases in some predefined step size.

Now recall the image registration problem over some transformation space τ with $D = MS$. Given by

$$\underset{\varphi \in \tau}{\text{minimize}} \quad MS(A, \varphi(B)) = \frac{1}{mn} \sum_{x,y} e_{xy}(\varphi(B)_{xy})$$

A Convex Image Registration Problem

Now recall the image registration problem over some transformation space τ with $D = MS$. Given by

$$\underset{\varphi \in \tau}{\text{minimize}} \quad MS(A, \varphi(B)) = \frac{1}{mn} \sum_{x,y} e_{xy}(\varphi(B)_{xy})$$

where

$$e_{xy}(\varphi(B)_{xy}) = (A_{xy} - \varphi(B)_{xy})^2$$

A Convex Image Registration Problem

Now recall the image registration problem over some transformation space τ with $D = MS$. Given by

$$\underset{\varphi \in \tau}{\text{minimize}} \quad MS(A, \varphi(B)) = \frac{1}{mn} \sum_{x,y} e_{xy}(\varphi(B)_{xy})$$

where

$$e_{xy}(\varphi(B)_{xy}) = (A_{xy} - \varphi(B)_{xy})^2$$

It is helpful to consider e_{xy} to be the cost function associated with (x, y), measuring the cost of the transformation φ for that pixel.

A Convex Image Registration Problem

Now recall the image registration problem over some transformation space τ with $D = MS$. Given by

$$\underset{\varphi \in \tau}{\text{minimize}} \quad MS(A, \varphi(B)) = \frac{1}{mn} \sum_{x,y} e_{xy}(\varphi(B)_{xy})$$

where

$$e_{xy}(\varphi(B)_{xy}) = (A_{xy} - \varphi(B)_{xy})^2$$

It is helpful to consider e_{xy} to be the cost function associated with (x, y), measuring the cost of the transformation φ for that pixel.

In order for this optimization problem to be convex we require:

1. τ to be a convex set.

Now recall the image registration problem over some transformation space τ with $D = MS$. Given by

$$\underset{\varphi \in \tau}{\text{minimize}} \quad MS(A, \varphi(B)) = \frac{1}{mn} \sum_{x,y} e_{xy}(\varphi(B)_{xy})$$

where

$$e_{xy}(\varphi(B)_{xy}) = (A_{xy} - \varphi(B)_{xy})^2$$

It is helpful to consider e_{xy} to be the cost function associated with (x, y), measuring the cost of the transformation φ for that pixel.

In order for this optimization problem to be convex we require:

1. τ to be a convex set.
2. MS to be a convex function.

A Convex Image Registration Problem

We construct a convex transformation space $\tau \subseteq \mathbb{R}^{2k}$ by parameterizing transformations in the following way:

- Define a collection functions in x and y, that form a set of basis vectors for a given image when evaluated at each pixel. *e.g.*

$$(1, x, y), \quad (1, x, y, xy, x^2, y^2), \quad (\gamma(\|(x, y) - (x_i, y_i)\|))_{i=1}^{k}$$

A Convex Image Registration Problem

We construct a convex transformation space $\tau \subseteq \mathbb{R}^{2k}$ by parameterizing transformations in the following way:

- Define a collection functions in x and y, that form a set of basis vectors for a given image when evaluated at each pixel. *e.g.*

$$(1, x, y), \quad (1, x, y, xy, x^2, y^2), \quad (\gamma(\|(x, y) - (x_i, y_i)\|))_{i=1}^{k}$$

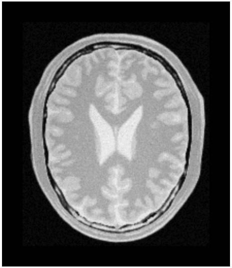
- Introduce variables $\mathbf{p_X}$ and $\mathbf{p_Y}$ in $\subseteq \mathbb{R}^k$, where k is the number of basis functions chosen, and define the displacement from (x, y) as a linear function of $\mathbf{p_X}$ and $\mathbf{p_Y}$. *e.g.*

$$(x, y) \mapsto \left(x + \mathbf{p_X}^T(1, x, y),\ \ y + \mathbf{p_Y}^T(1, x, y)\right)$$

$$(x, y) \mapsto \left(x + \sum_{i=1}^{k} p_{X_i}\gamma(r_i),\ \ y + \sum_{i=1}^{k} p_{Y_i}\gamma(r_i)\right)$$

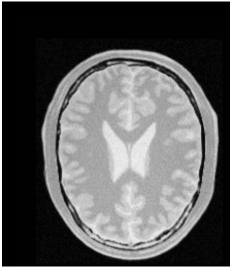
A Convex Image Registration Problem

To exemplify this, we model the translation below:



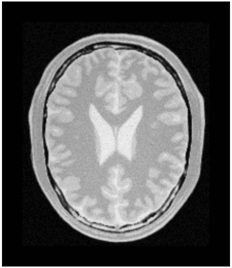
Base, B

$$\xrightarrow{\varphi(B)}$$



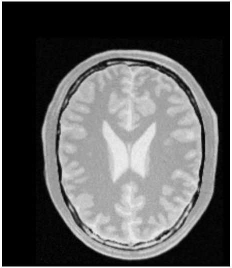
Target, A

To exemplify this, we model the translation below:



Base, B

$$\xrightarrow{\varphi(B)}$$

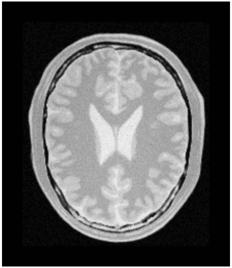


Target, A

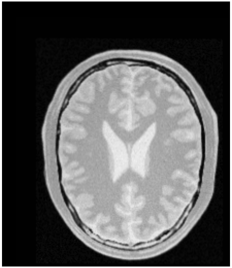
Using basis functions $(1, x, y)$ with

$$\mathbf{p_X} = (13, 0, 0) \text{ and } \mathbf{p_Y} = (17, 0, 0) \in \mathbb{R}^3.$$

To exemplify this, we model the translation below:



$$\xrightarrow{\varphi(B)}$$



Base, B

Target, A

Using basis functions $(1, x, y)$ with

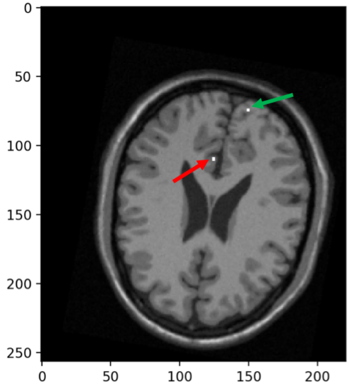
$$\mathbf{p_X} = (13, 0, 0) \text{ and } \mathbf{p_Y} = (17, 0, 0) \in \mathbb{R}^3.$$

i.e. for any pixel (x, y) we have

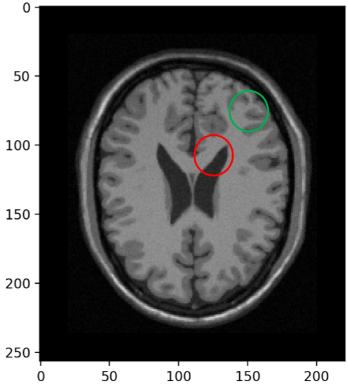
$$(x, y) \mapsto (x + 13, y + 17).$$

Now consider the behavior of a particular cost function e_{xy} of MS :

Now consider the behavior of a particular cost function e_{xy} of MS :

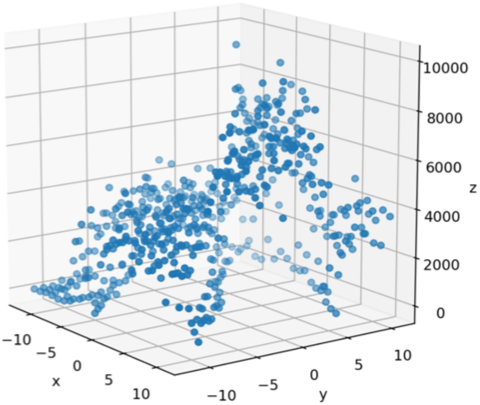


Target, A



Base, B

But clearly this is not the circumstance we find ourselves in.



A Convex Image Registration Problem

The proposed method by Taylor and Bhusnurmath is to then contruct a convex lower bound e'_{xy} of each individual cost function e_{xy} defined as follows:

A Convex Image Registration Problem

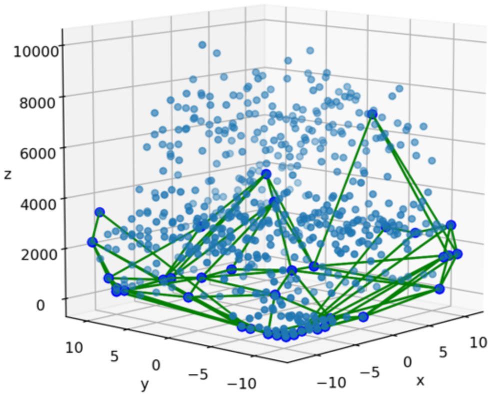
The proposed method by Taylor and Bhusnurmath is to then contruct a convex lower bound e'_{xy} of each individual cost function e_{xy} defined as follows:

- For each pixel (x, y) in the target image, replace error surface with its lower convex hull, or convex lower bound.

A Convex Image Registration Problem

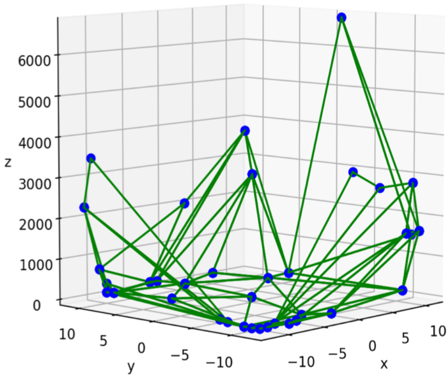
The proposed method by Taylor and Bhusnurmath is to then contruct a convex lower bound e'_{xy} of each individual cost function e_{xy} defined as follows:

- For each pixel (x, y) in the target image, replace error surface with its lower convex hull, or convex lower bound.

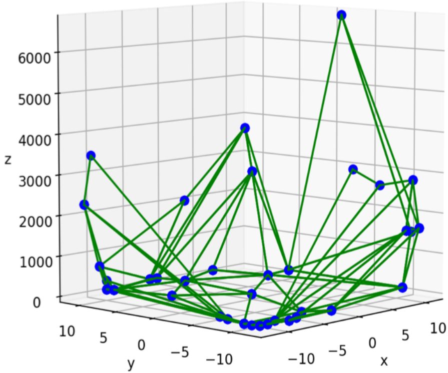


So for each pixel (x, y) we have an associated error surface:

So for each pixel (x, y) we have an associated error surface:



So for each pixel (x, y) we have an associated error surface:



$MS' = \frac{1}{mn} \sum_{x,y} e'_{xy}(\varphi(B)_{xy})$ gives a convex approximation to MS.

A Convex Image Registration Problem

The lower hull of a set of points in $\mathbb{R}^3$, as we have, can be described by the planar facets that define it.

A Convex Image Registration Problem

The lower hull of a set of points in $\mathbb{R}^3$, as we have, can be described by the planar facets that define it.

- A plane in $\mathbb{R}^3$ is defined by 4 coefficients a_1, a_2, a_3, b, so that a point (x, y, z) is in the plane if

$$a_1 x + a_2 y + a_3 z = b$$

A Convex Image Registration Problem

The lower hull of a set of points in $\mathbb{R}^3$, as we have, can be described by the planar facets that define it.

- A plane in $\mathbb{R}^3$ is defined by 4 coefficients a_1, a_2, a_3, b, so that a point (x, y, z) is in the plane if

$$a_1 x + a_2 y + a_3 z = b$$

- For each pixel (x, y) we have a some number, F_{xy}, of planes that define it, each with these four coefficients.

$$a_x^i(x,y)\tilde{x} + a_y^i(x,y)\tilde{y} + a_z^i(x,y)\tilde{z} = b_i(x,y), \quad i = 1, \ldots, F_{xy}.$$

The lower hull of a set of points in $\mathbb{R}^3$, as we have, can be described by the planar facets that define it.

- A plane in $\mathbb{R}^3$ is defined by 4 coefficients a_1, a_2, a_3, b, so that a point (x, y, z) is in the plane if

$$a_1 x + a_2 y + a_3 z = b$$

- For each pixel (x, y) we have a some number, F_{xy}, of planes that define it, each with these four coefficients.

  $$a_x^i(x,y)\tilde{x} + a_y^i(x,y)\tilde{y} + a_z^i(x,y)\tilde{z} = b_i(x,y), \quad i = 1, \ldots, F_{xy}.$$

  Here $\tilde{x}$ and $\tilde{y}$ represent displacements from x and y respectively, and $\tilde{z}$ is the point on the corresponding facet.

We introduce the notation $D(x, y, \mathbf{p_X})$ and $D(x, y, \mathbf{p_Y})$ to represent the displacements from x and y respectively, defined by $\mathbf{p_X}$ and $\mathbf{p_Y}$.

We introduce the notation $D(x, y, \mathbf{p_X})$ and $D(x, y, \mathbf{p_Y})$ to represent the displacements from x and y respectively, defined by $\mathbf{p_X}$ and $\mathbf{p_Y}$.

- The cost associated with any particular displacment should then be a value z such that the point

$$(D(x, y, \mathbf{p_X}),\ \ D(x, y, \mathbf{p_Y}),\ \ z)$$

lies on or above the convex error surface.

We introduce the notation $D(x, y, \mathbf{p_X})$ and $D(x, y, \mathbf{p_Y})$ to represent the displacements from x and y respectively, defined by $\mathbf{p_X}$ and $\mathbf{p_Y}$.

- The cost associated with any particular displacment should then be a value z such that the point

$$(D(x, y, \mathbf{p_X}),\ \ D(x, y, \mathbf{p_Y}),\ \ z)$$

lies on or above the convex error surface. To reflect this we introduce an auxiliary variable $z(x, y)$ for each pixel and require that, for $i = 1, \ldots, F_{xy}$,

$$a_x^i(x, y)D_x(x, y, \mathbf{p_X}) + a_y^i(x, y)D_y(x, y, \mathbf{p_Y}) - z(x, y) \geq b_i(x, y).$$

i.e. $z(x, y)$ must be on or above each planar facet of the new error surface, and then we have an associated cost $e'_{xy} = z(x, y)$.

A Convex Image Registration Problem

At this point we are ready to cast the registration problem as a convex optimization problem in the following way:

A Convex Image Registration Problem

At this point we are ready to cast the registration problem as a convex optimization problem in the following way:

$$\text{minimize}_{z,\mathbf{p_X},\mathbf{p_Y}} \sum_{x,y} z(x,y) = \mathbf{1}^T z$$

subject to

$$z(x,y) \geq a_x^i(x,y) D_x(x,y,\mathbf{p_X}) + a_y^i(x,y) D_y(x,y,\mathbf{p_Y}) - b_i(x,y)$$

for all i, for all x, y.

A Convex Image Registration Problem

At this point we are ready to cast the registration problem as a convex optimization problem in the following way:

$$\text{minimize}_{z,\mathbf{p_X},\mathbf{p_Y}} \sum_{x,y} z(x,y) = \mathbf{1}^T z$$

subject to

$$z(x,y) \geq a_x^i(x,y) D_x(x,y,\mathbf{p_X}) + a_y^i(x,y) D_y(x,y,\mathbf{p_Y}) - b_i(x,y)$$

for all i, for all x, y.

- This is a linear function of z with linear inequality constraints, of which there are $F = \sum_{x,y} F_{xy}$, one for each planar facet.

This is precisely the form of the problem we used the barrier method to solve.

We claim that, with blood, sweat and tears, this problem may be reformulated as

This is precisely the form of the problem we used the barrier method to solve.

We claim that, with blood, sweat and tears, this problem may be reformulated as

$$\begin{aligned}
&\text{minimize}_{\mathbf{p_X},\mathbf{p_Y},z}\ \ \mathbf{1}^T z \\
&\text{subject to} \\
&A_X C\mathbf{p_X} + A_Y C\mathbf{p_Y} - I_z z \preccurlyeq b.
\end{aligned}$$

It may be shown that, in solving each problem given by the barrier method, computing the descent direction $\delta_{\mathbf{p}}$, can be boiled down to solving the following equation:

$$(H_{\mathbf{p}} - H_z D^{-1} H_z^T)\delta_{\mathbf{p}} = g_{\mathbf{p}} - H_z D^{-1} g_z.$$

It may be shown that, in solving each problem given by the barrier method, computing the descent direction $\delta_{\mathbf{p}}$, can be boiled down to solving the following equation:

$$(H_{\mathbf{p}} - H_z D^{-1} H_z^T)\delta_{\mathbf{p}} = g_{\mathbf{p}} - H_z D^{-1} g_z.$$

The size of this system depends only on the number of parameters used to describe the transformation. *i.e.* $\mathbf{p_X}, \mathbf{p_Y} \in \mathbb{R}^k$, gives a system of $2k$ linear equations.

It may be shown that, in solving each problem given by the barrier method, computing the descent direction $\delta_{\mathbf{p}}$, can be boiled down to solving the following equation:

$$(H_{\mathbf{p}} - H_z D^{-1} H_z^T)\delta_{\mathbf{p}} = g_{\mathbf{p}} - H_z D^{-1} g_z.$$

The size of this system depends only on the number of parameters used to describe the transformation. *i.e.* $\mathbf{p_X}, \mathbf{p_Y} \in \mathbb{R}^k$, gives a system of $2k$ linear equations.

- *e.g.*, an affine transformation described by the basis functions $(1, x, y)$ requires $\mathbf{p_X}, \mathbf{p_Y} \in \mathbb{R}^3$, so that the above is as system of 6 equations.

So we have successfully constructed a convex problem related to the image registration problem, but is it at all meaningful or useful?

So we have successfully constructed a convex problem related to the image registration problem, but is it at all meaningful or useful?

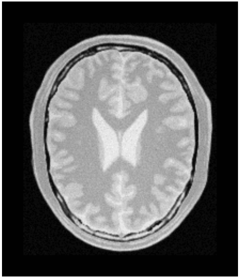
- We saw that the cost associated with a particular pixel may be severely underestimated via our construction. Isn't that bad?

Results

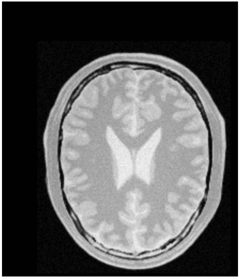
So we have successfully constructed a convex problem related to the image registration problem, but is it at all meaningful or useful?

- We saw that the cost associated with a particular pixel may be severely underestimated via our construction. Isn't that bad?
- These are images, after all, how feasible is it to construct this approximation for each pixel?

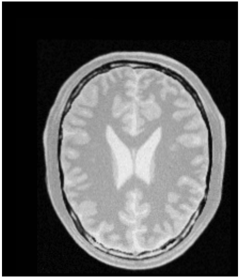
Results (Rigid)



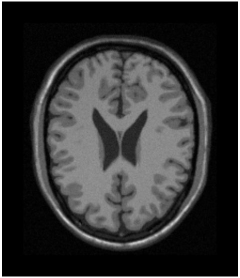
Base, B



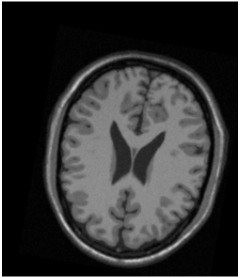
Target, A



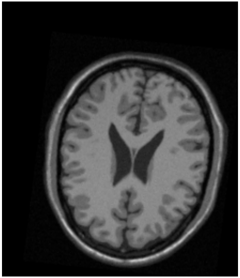
Registered, $\Phi(B)$



Base, B

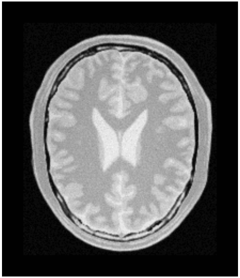


Target, A

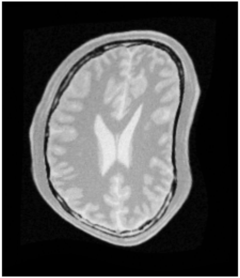


Registered, $\Phi(B)$

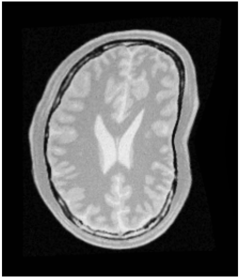
Results (Non-Rigid)



Base, B

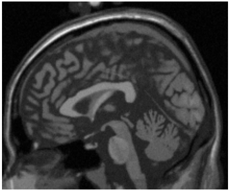


Target, A

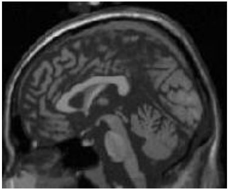


Registered, $\Phi(B)$

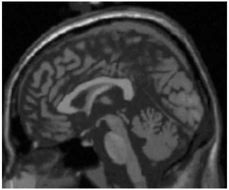
Results (Non-Rigid)



Base, B

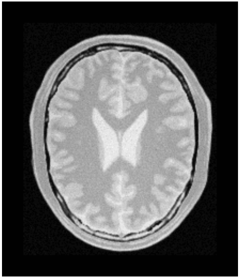


Target, A

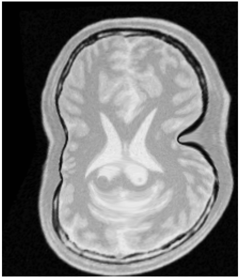


Registered, $\Phi(B)$

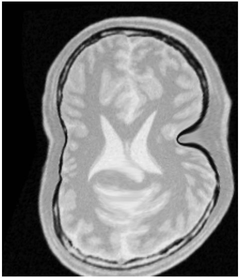
Results (Non-Rigid)



Base, B

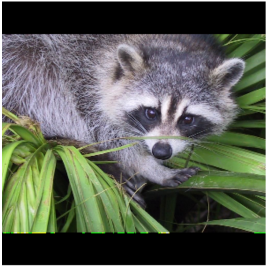


Target, A

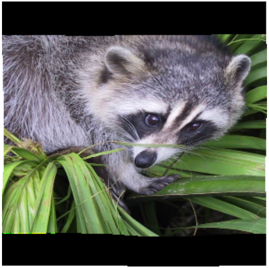


Registered, $\Phi(B)$

Results (Non-Rigid)

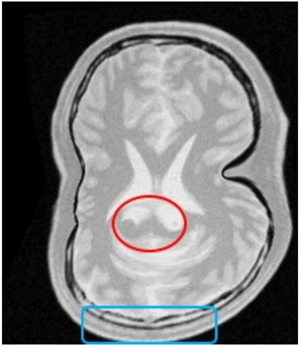


Base, B

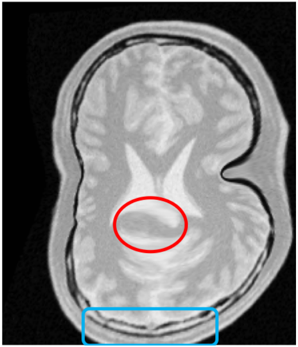


Target, A

Registered, $\Phi(B)$

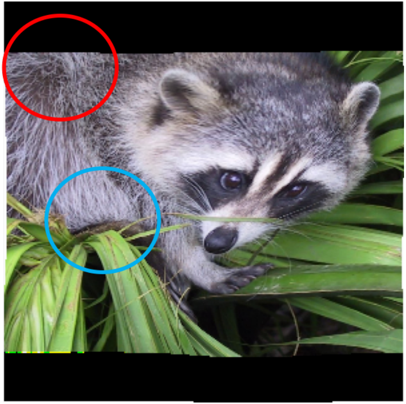


Target, A

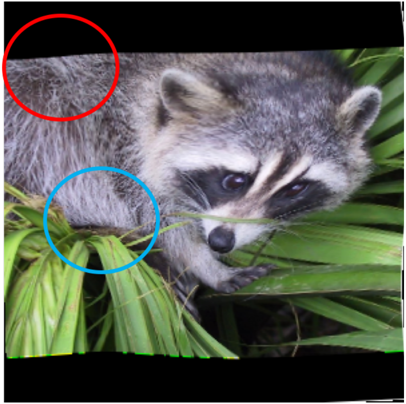


Registered, $\Phi(B)$

Results



Target, A



Registered, $\Phi(B)$

Limitations and Future Work

- The convex approximation is dependent upon some window size w, directly impacting each planar facet.

Limitations and Future Work

- The convex approximation is dependent upon some window size w, directly impacting each planar facet.
- Sensitivity to noise.

Limitations and Future Work

- The convex approximation is dependent upon some window size w, directly impacting each planar facet.
- Sensitivity to noise.
- A multi-modal convex registration algorithm

Thank You

Thank you to my advisor, the entire math department, my family and my friends.

Thank You

Thank you to my advisor, the entire math department, my family and my friends.

Questions?