2022 UW-Madison Analysis REU

CONVEXITY

1 Motivation

- General Function

  👎👎 very hard or impossible to find the global minimum; best to hope for is to find local minimum

- General Convex Function

  👍 local minimum is guaranteed to be global minimum

  👎 there might be no minimum or multiple minima

- Strictly Convex Function *(very important concept in the CDR paper)*

  👍 local minimum is guaranteed to be the UNIQUE global minimum

  👎 again, it may not exist

- Strongly Convex Function *(not used in the CDR paper though)*

  👍👍 there ALWAYS exists a UNIQUE global minimum

2 Basic Definitions

Throughout this document, let $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}^d$, be a twice continuously differentiable function and $\mathbf{x}_0$ be an interior point of D.

2.1 Local & Global Minimizer

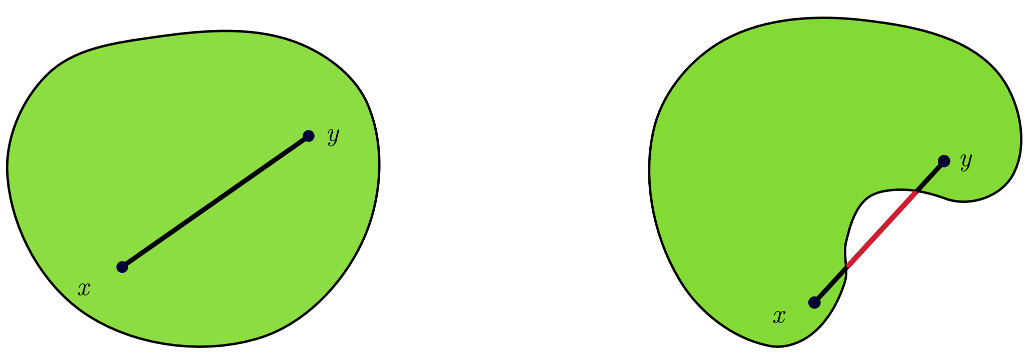
$\mathbf{x}_0$ is a *local* minimizer if there exists $\delta > 0$ such that $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ for all $\mathbf{x}$ in $B_\delta(\mathbf{x}_0)\backslash\{\mathbf{x}_0\}$.

$\mathbf{x}_0$ is a *global* minimizer if $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ for all $\mathbf{x}$ in D.

$\mathbf{x}_0$ is a strict *local* or *global* minimizer if the above inequality is strict, respectively.

2.2 Stationary Point

$\mathbf{x}_0$ is a stationary point of f if $\nabla f(\mathbf{x}_0) = 0$.



(a) a convex set in $\mathbb{R}^2$ (b) a non-convex set in $\mathbb{R}^2$

Figure 1: Convex Set vs Non-convex Set

2.3 Convex Set

A set $D \in \mathbb{R}^d$ is convex if for all $\mathbf{x}, \mathbf{y} \in D$ and all $\lambda \in [0, 1]$,

$$(1-\lambda)\,\mathbf{x} + \lambda\,\mathbf{y} \in D.$$

2.4 Convex Function

f is a convex function if for all $\mathbf{x}, \mathbf{y} \in D$ and all $\lambda \in [0, 1]$,

$$f((1-\lambda)\,\mathbf{x} + \lambda\,\mathbf{y}) \leq (1-\lambda)f(\mathbf{x}) + (\mathbf{y}).$$

Equivalently, this condition can be written as

$$\Big(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\Big)^T (\mathbf{x} - \mathbf{y}) \geq 0.$$

This will be proved after we have the First-order Convexity Condition (4.1.1).

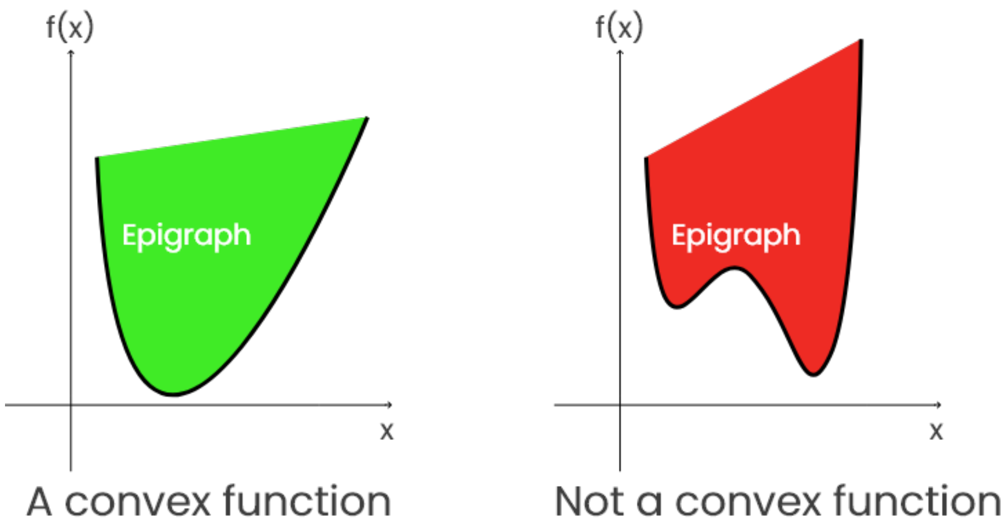


Figure 2: Convex Function vs Non-convex Function

2.5 Hessian

The matrix of $f(\mathbf{x}_0)$'s second derivatives is called the Hessian and is denoted by

$$\mathbf{H}_f\,(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_d \partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_1 \partial x_d} & \cdots & \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_d^2} \end{pmatrix}.$$

2.6 Definite Matrix

Let $M \in \mathbb{R}^{d \times d}$ be a symmetric matrix and $\mathbf{v} \in \mathbb{R}^d$, we say M is positive semidefinite if $\mathbf{v}^T\,M\,\mathbf{v} \geq 0$. We say M is positive definite if the inequality is strict.

3 Some Important Theorems

3.1 General Optimality Conditions

3.1.1 First-order Necessary Condition

$\mathbf{x}_0$ is a local minimizer of $f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}$

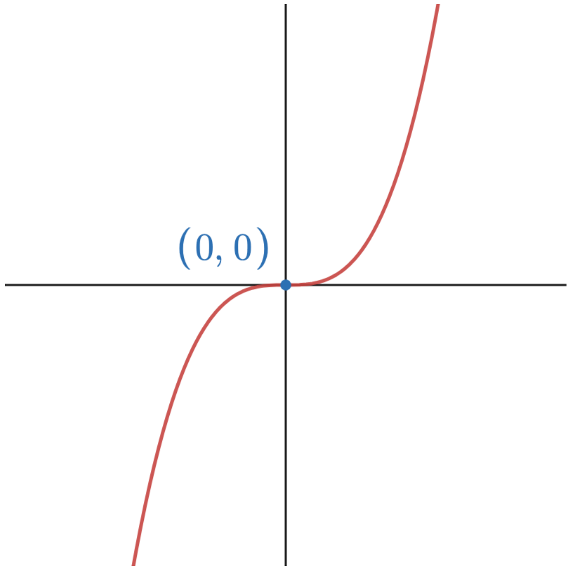


Figure 3: zero is a stationary point of $f(x) = x^3$, but it's not a local minimizer

3.1.2 Second-order Necessary Condition

$\mathbf{x}_0$ is a local minimizer of $f \implies \nabla f(\mathbf{x}_0) = \mathbf{0}$ and $H_f(\mathbf{x}_0)$ is positive semidefinite

3.1.3 Second-order Sufficient Condition

$\nabla f(\mathbf{x}_0) = \mathbf{0}$ and $H_f(\mathbf{x}_0)$ is positive definite $\implies \mathbf{x}_0$ is a strict local minimizer

3.2 Mean Value Theorem

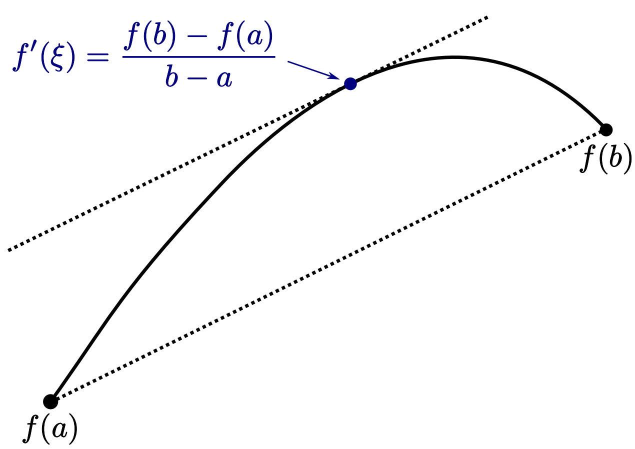


Figure 4: Demo of the Mean Value Theorem when $D \subseteq \mathbb{R}$

For any $\mathbf{x} \in B_\delta(\mathbf{x}_0)$, there exists some $\varepsilon \in (0,1)$ such that

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0 + \varepsilon(\mathbf{x}-\mathbf{x}_0))^T(\mathbf{x}-\mathbf{x}_0) \\ \frac{f(\mathbf{x}) - f(\mathbf{x}_0)}{(\mathbf{x}-\mathbf{x}_0)} &= \nabla f(\mathbf{x}_0 + \varepsilon(\mathbf{x}-\mathbf{x}_0))^T, \end{aligned}$$

3.3 Second-order Taylor Expansion

Let $\mathbf{x} \in B_\delta(\mathbf{x}_0)$ for some $\delta > 0$, then there exists $\varepsilon > 0$ such that

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x}-\mathbf{x}_0) + \frac{1}{2}(\mathbf{x}-\mathbf{x}_0)^T H_f(\mathbf{x}_0 + \varepsilon(\mathbf{x}-\mathbf{x}_0))(\mathbf{x}-\mathbf{x}_0)$$

4 Selected Topics in Convex Function

4.1 Convexity Condition

4.1.1 First-order Convexity Condition

f is convex if and only if $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in D$.

(First Direction) $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \implies f$ is convex:

Let $\mathbf{z} = (1-\lambda)\,\mathbf{x} + \lambda\,\mathbf{y}$, where $\lambda \in [0,1]$.

$$f(\mathbf{x}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x}-\mathbf{z}) \tag{1}$$
$$f(\mathbf{y}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{y}-\mathbf{z}) \tag{2}$$

$(1-\lambda)\cdot(1) + \lambda\cdot(2)$ would give

$$\begin{aligned} (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) &\geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T\Big((1-\lambda)(\mathbf{x}-\mathbf{z}) + \lambda(\mathbf{y}-\mathbf{z})\Big) \\ &= f(\mathbf{z}) + \nabla f(\mathbf{z})^T(\mathbf{x} - \lambda\,\mathbf{x} - \mathbf{z} + \cancel{\lambda\mathbf{z}} + \lambda\,\mathbf{y} - \cancel{\lambda\mathbf{z}}) \\ &= f(\mathbf{z}) + \nabla f(\mathbf{z})^T\Big(\cancelto{0}{(1-\lambda)\,\mathbf{x} + \lambda\,\mathbf{y} - \mathbf{z}}\Big) \\ &= f\Big((1-\lambda)\,\mathbf{x} + \lambda\,\mathbf{y}\Big) \end{aligned}$$

□

(Second Direction) f is convex $\implies f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x})$:

By the Mean Value Theorem, since $\mathbf{z} = \mathbf{x} + \lambda(\mathbf{y}-\mathbf{x}) \in B(\mathbf{x})$, there exists ε such that

$$f(\mathbf{z}) = f(\mathbf{x} + \lambda(\mathbf{y}-\mathbf{x})) = f(\mathbf{x}) + \nabla f(\mathbf{x} + \varepsilon(\mathbf{z}-\mathbf{x}))^T(\mathbf{z}-\mathbf{x}). \tag{3}$$

By convexity,

$$f(\mathbf{z}) = f\Big((1-\lambda)\,\mathbf{x} + \lambda\,\mathbf{y}\Big) \leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}). \tag{4}$$

Combine (3) and (4),

$$\begin{aligned} f(\mathbf{x}) + \nabla f(\mathbf{x} + \varepsilon(\mathbf{z}-\mathbf{x}))^T(\mathbf{z}-\mathbf{x}) &\leq (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) \\ \cancel{\lambda} f(\mathbf{x}) + \nabla f(\mathbf{x} + \varepsilon(\mathbf{z}-\mathbf{x}))^T \cancel{\lambda}(\mathbf{y}-\mathbf{x}) &\leq \cancel{\lambda} f(\mathbf{y}) \\ f(\mathbf{x}) + \nabla f(\mathbf{x} + \varepsilon\lambda(\mathbf{y}-\mathbf{x}))^T(\mathbf{y}-\mathbf{x}) &\leq f(\mathbf{y}). \end{aligned}$$

Taking $\lambda \to 0$,

$$f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}). \qquad \square$$

Now that we've proved the First-order Convexity Condition, an immediate implication is that, as mentioned earlier, the definition of convex function can be rewritten as

$$\Big(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\Big)^T (\mathbf{x} - \mathbf{y}) \geq 0,$$

for all $\mathbf{x}, \mathbf{y} \in D$.

Proof: By the First-order Convexity Condition,

$$\begin{aligned}
\Big(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\Big)^T (\mathbf{x} - \mathbf{y}) &= \nabla f(\mathbf{x})^T(\mathbf{x} - \mathbf{y}) - \nabla f(\mathbf{y})^T(\mathbf{x} - \mathbf{y}) \\
&\geq f(\mathbf{x}) - f(\mathbf{y}) + f(\mathbf{y}) - f(\mathbf{x}) \\
&= 0.
\end{aligned} \qquad \square$$

4.1.2 Second-Order Convexity Condition

f is convex if and only if $H_f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in D$ (proof omitted).

4.2 Property of Convex Function

4.2.1 Convex Function's Local Minimizer is also Global Minimizer

We prove by contradiction. Suppose $\mathbf{x}_0$ is a local minimizer but not a global minimizer of f, then there exists $\mathbf{y} \in D$ such that $f(\mathbf{y}) < f(\mathbf{x}_0)$. Since f is convex, for any $\lambda \in [0, 1]$,

$$\begin{aligned}
f\Big((1-\lambda)\,\mathbf{x}_0 + \lambda\,\mathbf{y}\Big) &\leq (1-\lambda) f(\mathbf{x}_0) + \lambda f(\mathbf{y}) \\
&< f(\mathbf{x}_0),
\end{aligned}$$

which implies in every open ball around $\mathbf{x}_0$ there exists points on the line segment connecting $\mathbf{x}_0$ and $\mathbf{y}$ that take smaller values than $f(\mathbf{x}_0)$. This contradicts the assumption that $\mathbf{x}_0$ is a local minimizer. $\square$

Conversely, a concave function's local maximizer is also global maximizer.

4.2.2 Stationary Point of Convex Function is Global Minimizer

It immediately follows that if f is convex, then a stationary point of f must be a global minimizer.

4.3 Strictly Convex Function

f is strictly convex if the inequality in 2.4 is strict.

4.3.1 A Strictly Convex Function's Global Minimizer is Unique (if it exists)

We prove by contradiction and assume f is strictly convex but it has 2 global minimizers $\mathbf{y}, \mathbf{z} \in D$. Then $f(\mathbf{y}) = f(\mathbf{z})$ and $f(\mathbf{y}), f(\mathbf{z}) \geq f(\mathbf{x})$ for all $\mathbf{x} \in D$. For any $\lambda \in [0, 1]$,

$$\begin{aligned}
f\Big((1-\lambda)\,\mathbf{y} + \lambda\,\mathbf{z}\Big) &< (1-\lambda) f(\mathbf{y}) + \lambda f(\mathbf{z}) \\
&= f(\mathbf{y}),
\end{aligned}$$

which indicates there exists points that take strictly smaller values than $\mathbf{y}$. This contradicts the assumption that $\mathbf{y}$ is a global minimizer. $\square$

4.4 Strongly Convex Function

f is m-strongly convex if $H_f(\mathbf{x}) \succeq mI_{d\times d}$ for all $\mathbf{x} \in D$, where $m > 0$.

The notations '$\succeq$' and '$\preceq$' are used to denote that

$$A \succeq 0 \iff A \text{ is positive semidefinite,}$$
$$A \preceq 0 \iff A \text{ is negative semidefinite,}$$

where $A \in \mathbb{R}^{d\times d}$ is a matrix. These inequalities could be expressed in terms of '$\geq$' and '$\leq$' in the following way:

$$A \succeq B \iff \mathbf{v}^T A\,\mathbf{v} \geq \mathbf{v}^T B\,\mathbf{v},$$
$$A \preceq B \iff \mathbf{v}^T A\,\mathbf{v} \leq \mathbf{v}^T B\,\mathbf{v},$$

where $\mathbf{v} \in \mathbb{R}^d$. Therefore, the condition of m-strongly convex function can be rewritten as, for all $\mathbf{x} \in D$,

$$\begin{aligned} \mathbf{x}^T H_f(\mathbf{x})\,\mathbf{x} &\geq \mathbf{x}^T mI_{d\times d}\,\mathbf{x} \\ &= m\mathbf{x}^T I_{d\times d}\,\mathbf{x} \\ &= m\|\,\mathbf{x}\,\|^2. \end{aligned} \tag{5}$$

Equivalently, this condition can be written as yet another form:

$$\Big(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\Big)^T (\mathbf{x} - \mathbf{y}) \geq m\|\,\mathbf{x} - \mathbf{y}\,\|^2, \tag{6}$$

for all $\mathbf{x}, \mathbf{y} \in D$.

4.4.1 Quadratic Bound for Strongly Convex Function

Suppose f is m-strongly convex. Let $\mathbf{y} \in B_\delta(\mathbf{x})$ for $\delta > 0$ and write $f(\mathbf{y})$ in the second-order Taylor expansion form:

$$\begin{aligned} f(\mathbf{y}) &= f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T H_f(\mathbf{x} + \varepsilon(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \\ &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) + \frac{m}{2}\|\,\mathbf{y} - \mathbf{x}\,\|^2. && \text{by (5)} \end{aligned}$$

This quadratic bound holds for all $\mathbf{x}, \mathbf{y} \in D$. □

Note that this quadratic bound implies that a strongly convex function must also be strictly convex.

4.4.2 Stationary Point of Strongly Convex Function is Unique Global Minimizer

Proof: Suppose $\nabla f(\mathbf{x}_0) = 0$, by the quadratic bound (4.4.1),

$$\begin{aligned} f(\mathbf{y}) &\geq f(\mathbf{x}_0) + \cancel{\nabla f(\mathbf{x}_0)^T(\mathbf{y} - \mathbf{x}_0)} + \frac{m}{2}\|\,\mathbf{y} - \mathbf{x}_0\,\|^2 \\ &= f(\mathbf{x}_0) + \frac{m}{2}\|\,\mathbf{y} - \mathbf{x}_0\,\|^2 \\ &> f(\mathbf{x}_0), \end{aligned}$$

for all $\mathbf{y} \neq \mathbf{x}_0 \in D$. Thus the stationary point $\mathbf{x}_0$ of f is a unique global minimizer. □

4.4.3 Existence of Strongly Convex Function's Global Minimizer

First, we prove that if f is m-strongly convex, then it's coercive, i.e.,

$$\lim_{\|\,\mathbf{y}\,\|\to\infty} f(\mathbf{y}) = +\infty \tag{7}$$

Proof: Let $g(\mathbf{x}) = f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2$ for any $C \le m$, then for any $\mathbf{x}, \mathbf{y} \in D$,

$$\begin{aligned}
\Big(\nabla g(\mathbf{y}) - \nabla g(\mathbf{x})\Big)(\mathbf{y}-\mathbf{x}) &= \Big(\nabla\big(f(\mathbf{y}) - \frac{C}{2}\|\,\mathbf{y}\,\|^2\big) - \nabla\big(f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2\big)\Big)^T(\mathbf{y}-\mathbf{x}) \\
&= \Big(\nabla f(\mathbf{y}) - C\,\mathbf{y} - \nabla f(\mathbf{x}) - C\,\mathbf{x}\Big)^T(\mathbf{y}-\mathbf{x}) \\
&= \left\{\Big(\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})\Big)^T(\mathbf{y}-\mathbf{x})\right\} - (C\,\mathbf{y} - C\,\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \\
&\ge \Big\{m\|\,\mathbf{y}-\mathbf{x}\,\|^2\Big\} - C\|\,\mathbf{y}-\mathbf{x}\,\|^2 && \text{by (6)} \\
&= (m - C)\|\,\mathbf{y}-\mathbf{x}\,\|^2 \\
&\ge 0
\end{aligned}$$

Therefore, by the alternative definition of convex function (2.4), g is convex. Since g is convex, by first-order convex condition (4.1), for all $\mathbf{x}, \mathbf{y} \in D$,

$$\begin{aligned}
g(\mathbf{y}) &\ge g(\mathbf{x}) + \nabla g(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \\
f(\mathbf{y}) - \frac{C}{2}\|\,\mathbf{y}\,\|^2 &\ge f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2 + \nabla\Big(f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2\Big)^T(\mathbf{y}-\mathbf{x}) \\
f(\mathbf{y}) &\ge \frac{C}{2}\|\,\mathbf{y}\,\|^2 + f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2 + \nabla\Big(f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2\Big)^T\mathbf{y} - \Big(f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2\Big)^T\mathbf{x} \\
&= \left\{\frac{C}{2}\|\,\mathbf{y}\,\|^2 + \nabla\Big(f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2\Big)^T\mathbf{y}\right\} + P_{\mathbf{x}} && \text{where } P_{\mathbf{x}} \text{ only depends on } \mathbf{x} \\
&\ge \left\{\frac{C}{2}\|\,\mathbf{y}\,\|^2 - \left\|\nabla\Big(f(\mathbf{x}) - \frac{C}{2}\|\,\mathbf{x}\,\|^2\Big)\right\|\|\,\mathbf{y}\,\|\right\} + P_{\mathbf{x}} && (8)
\end{aligned}$$

Taking $\|\,\mathbf{y}\,\| \to \infty$, (8)$\to \infty$. $\square$

Second, we prove that for any $\mathbf{x}_0 \in D$, the set

$$C_{\mathbf{x}_0} = \{\mathbf{x} \in D : f(\mathbf{x}) \le f(\mathbf{x}_0)\}$$

is compact, i.e., closed and bounded.

Proof: For any $\mathbf{x} \in C^c_{\mathbf{x}_0}$, $f(\mathbf{x}) > f(\mathbf{x}_0)$. Since f is smooth, for any $\mathbf{v} \in D$, there exists $\alpha > 0$ small enough such that $f(\mathbf{x} \pm \alpha\mathbf{v}) > f(\mathbf{x}_0)$, i.e., there exists $\delta > 0$ small enough such that $B_\delta(\mathbf{x}) \subset C^c_{\mathbf{x}_0}$. Therefore, $C^c_{\mathbf{x}_0}$ is open and thus $C_{\mathbf{x}_0}$ is closed.

By contradiction, we assume that there exists $\mathbf{y} \in C_{\mathbf{x}_0}$ such that $\|\,\mathbf{y}\,\|$ is unboundedly large. Then, by (7), $f(\mathbf{y}) \to +\infty$. This is a contradiction because $f(\mathbf{y}) \le f(\mathbf{x}_0)$. Therefore, $C_{\mathbf{x}_0}$ is bounded. $\square$

Finally, if pick any $\mathbf{x}_0 \in D$, we know that $C_{\mathbf{x}_0}$ is non-empty because it at least contains $\mathbf{x}_0$. By the Extreme Value Theorem, f attains its minimum on the compact set $C_{\mathbf{x}_0}$, which contains the global minimizer of f.

5 CDR Main Paper Optimization Problem (on p.224)

5.1 The Problem

The objective function is

$$O(\mathbf{v}) = \underbrace{\sum_{i \in V} \|\mathbf{v}_i\|^2}_{A(\mathbf{v})} + \underbrace{\sum_{\{i,j\} \in S} \frac{1}{(\mathbf{v}_i - \mathbf{v}_j) \cdot (\mathbf{p}_i - \mathbf{p}_j) - \|\mathbf{p}_j - \mathbf{p}_i\|}}_{B(\mathbf{v})}.$$

The constraints are

$$\begin{aligned}
&(\mathbf{v}_j - \mathbf{v}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) > \|\mathbf{p}_j - \mathbf{p}_i\|, \quad \text{for} \quad \{i,j\} \in S, \\
&(\mathbf{v}_j - \mathbf{v}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) = 0, \quad \text{for} \quad \{i,j\} \in B, \\
&\mathbf{v}_1 = \mathbf{v}_2 = 0.
\end{aligned}$$

5.2 Three Claims

Let $R = O(\mathbf{v})$ for some optimal feasible solution $\mathbf{v} \in \mathbb{R}^{2|V|}$ to the above optimization problem.

Claim 1: There exists some ball $D(\mathbf{0}, r)$ (this is the closed ball and includes its boundary) such that for all vectors $\mathbf{w}$ with $\mathbf{w}$ not in $D(\mathbf{0}, r)$, we have $A(\mathbf{w}) > 10R$. Therefore, $O(\mathbf{w}) > 10R$ whenever $\mathbf{w}$ not in $D(\mathbf{0}, r)$.

Proof: The closed ball is defined as $D(\mathbf{0}, r) = \{\mathbf{x} \in \mathbb{R}^{2|V|} : \|\,\mathbf{x}\,\| \leq r\}$. Suppose $\mathbf{w} \notin D(\mathbf{0}, r)$, then $\|\,\mathbf{w}\,\| > r$.

$$\begin{aligned}
A(\mathbf{w}) &= \sum_{i \in V} \|\,\underset{i}{\mathbf{w}}\,\|^2 \\
&= \sum_{i \in V} (w_i^x)^2 + (w_i^y)^2 \\
&= \|\,\mathbf{w}\,\| \\
&> r
\end{aligned}$$

Therefore, $O(\mathbf{w}) = A(\mathbf{w}) + B(\mathbf{w}) > r$. Let $r = 10R$, then $O(\mathbf{w}) > 10R$. $\square$

Claim 2: There exists some small value epsilon so that whenever $\mathbf{w}$ is within epsilon of the boundary of the region of feasibility, then $B(\mathbf{w}) > 10R$. Therefore, $O(\mathbf{w}) > 10R$ whenever $\mathbf{w}$ is within epsilon of the boundary of the region of feasibility.

Proof: Define the region within epsilon of the boundary of the region of feasibility as

$$K = \left\{ \mathbf{w} \in \mathbb{R}^{2|V|} : \left\{ \begin{array}{l} \|\mathbf{p}_j - \mathbf{p}_i\| < (\mathbf{w}_j - \mathbf{w}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) < \|\mathbf{p}_j - \mathbf{p}_i\| + \varepsilon, \quad \text{for} \quad \{i,j\} \in S \\ (\mathbf{w}_j - \mathbf{w}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) = 0, \quad \text{for} \quad \{i,j\} \in B \\ \mathbf{w}_1 = \mathbf{w}_2 = 0 \end{array} \right. \right\}.$$

Then for all $\mathbf{w} \in K$,

$$\begin{aligned}
B(\mathbf{w}) &= \sum_{\{i,j\} \in S} \frac{1}{(\mathbf{v}_i - \mathbf{v}_j) \cdot (\mathbf{p}_i - \mathbf{p}_j) - \|\mathbf{p}_j - \mathbf{p}_i\|} \\
&> \sum_{\{i,j\} \in S} \frac{1}{\varepsilon} \\
&= \frac{|S|}{\varepsilon}
\end{aligned}$$

Let $\varepsilon = \frac{|S|}{10R}$, then $B(\mathbf{w}) > 10R$. $\square$

Claim 3: The intersection of the region of feasibility, the closed ball $D(0, r)$, and all points that are not closer than epsilon to the boundary is a closed and bounded set in euclidean space.

Proof: Define the region of feasibility F as

$$F = \left\{ \mathbf{x} \in \mathbb{R}^{2|V|} : \begin{cases} (\mathbf{x}_j - \mathbf{x}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) < \|\mathbf{p}_j - \mathbf{p}_i\|, & \text{for} \quad \{i, j\} \in S \\ (\mathbf{x}_j - \mathbf{x}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) = 0, & \text{for} \quad \{i, j\} \in B \\ \mathbf{x}_1 = \mathbf{x}_2 = 0 \end{cases} \right\}.$$

Then the intersection in claim 3 is

$$N = D(\mathbf{0}, r) \cap F \backslash K. \tag{9}$$

We know that $D(\mathbf{0}, r)$ is compact, i.e., closed and bounded, since it's a closed ball. Now let $\mathbf{w} \in F \backslash K$, then

$$\begin{cases} (\mathbf{w}_j - \mathbf{w}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) \geq \|\mathbf{p}_j - \mathbf{p}_i\| + \varepsilon, & \text{for} \quad \{i, j\} \in S, \\ (\mathbf{w}_j - \mathbf{w}_i) \cdot (\mathbf{p}_j - \mathbf{p}_i) = 0, & \text{for} \quad \{i, j\} \in B, \\ \mathbf{w}_1 = \mathbf{w}_2 = 0, \end{cases}$$

which implies $F \backslash K$ is closed. Thus the intersection N is closed. And since $D(\mathbf{0}, r)$ is bounded, N must also be bounded. Therefore, N is compact. $\square$

5.3 Existence of Unique Global Minimizer

Through the 3 claims above, we've narrowed down the 'optimal region of feasibility' to a smaller set N in (9), because anything that is not in N won't attain the minimum. Now since N is compact, by the Extreme Value Theorem, O attains the minimum over N, i.e., there exists $\mathbf{v}^* \in N$ such that $O(\mathbf{v}^*) \leq O(\mathbf{x})$ for all $\mathbf{x} \in N$. And since O is strictly convex, the global minimizer $\mathbf{v}^*$ is unique.

Acknowledgement

The structure of this document follows Professor Jack Burkart's instructions. Lots of the material and proof ideas in this document are adapted from UW-Madison Math 535 lecture notes, Mathematics Stack Exchange, Wikipedia, and Jack.

A link to the CDR paper.