Regularity of Convex Functions on $\mathbb{R}^d$

Or : How geometry can help analysis !
Or^2 *: How geometry can create regularity !*

Definitions and Basics

Definition

A function $\phi : \mathbb{R}^d \to \mathbb{R}$ is **convex** if for all $x, y \in \mathbb{R}^d$ and for all $t \in [0, 1]$:

$$\phi(tx + (1-t)y) \leq t\phi(x) + (1-t)\phi(y)$$

Examples: $y = x^2, \quad y = |x|$

Remark

A function ϕ is convex if and only if its *epigraph* (the region above its graph) is a convex subset of $\mathbb{R} \times \mathbb{R}^d$.

Part I

Continuity of Convex Functions

Continuity Theorems

Main Theorem

Any convex function $\phi : \mathbb{R}^d \to \mathbb{R}$ is continuous.

To prove this, we rely on two fundamental lemmas:

Lemma 1: Local Boundedness

ϕ is bounded on any closed ball B_R for all $R > 0$.

Lemma 2: Local Lipschitz Continuity

ϕ is locally Lipschitz continuous on $\mathbb{R}^d$.
That is: $\forall R > 0, \ \exists L > 0, \ \forall x, y \in B_R \quad |\phi(x) - \phi(y)| \leq L\|x - y\|$

Proof of Lemma 1 (Local Boundedness) - Part 1/2

Upper Bound

Fix $R > 0$ and let $x \in B_R$. Notice that B_R is contained in the hypercube $C_R = [-R, R]^d$.

Denote by $(v_i)_{i=1}^{2^d}$ the vertices of the hypercube. We can write x as a convex combination of these vertices:

$$x = \sum_{i=1}^{2^d} \lambda_i v_i \quad \text{with} \quad \lambda_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{2^d} \lambda_i = 1$$

By convexity:

$$\phi(x) \leq \sum_{i=1}^{2^d} \lambda_i \phi(v_i) \leq \max_{i=1}^{2^d} \phi(v_i) := M$$

Proof of Lemma 1 (Local Boundedness) - Part 2/2

Lower Bound

Using central symmetry at the origin:

$$0 = \frac{1}{2}x + \frac{1}{2}(-x)$$

By convexity:

$$\phi(0) \leq \frac{1}{2}\phi(x) + \frac{1}{2}\phi(-x)$$

Rearranging gives:

$$\phi(x) \geq 2\phi(0) - \phi(-x)$$

Since $-x \in B_R$, we know from the upper bound that $\phi(-x) \leq M$, yielding:

$$\phi(x) \geq 2\phi(0) - M := m$$

Conclusion: $m \leq \phi(x) \leq M$ for all $x \in B_R$. □

Proof of Lemma 2 (Local Lipschitz) - Part 1/2

The Secant Argument

Fix $R > 0$ and let $x, y \in B_{R/2}$.
Define z as the intersection of the ray $[x, y)$ with the boundary ∂B_R.
We can express y as a convex combination of x and z:

$$y = (1-t)x + tz \quad \text{where} \quad t = \frac{\|y-x\|}{\|z-x\|}$$

By convexity:

$$\phi(y) \leq (1-t)\phi(x) + t\phi(z)$$

Rearranging:

$$\phi(y) - \phi(x) \leq \frac{\|y-x\|}{\|z-x\|}(\phi(z) - \phi(x))$$

Proof of Lemma 2 (Local Lipschitz) - Part 2/2

Geometric Bounds

Since $x \in B_{R/2}$ and $z \in \partial B_R$, the distance $\|z - x\| \geq \frac{R}{2}$.
Furthermore, we know $m \leq \phi \leq M$ on the ball B_R.
Thus:

$$\phi(y) - \phi(x) \leq \frac{2}{R}(M - m)\|y - x\|$$

By a symmetric argument swapping the roles of x and y, we obtain the same inequality for $\phi(x) - \phi(y)$.

Conclusion: $|\phi(y) - \phi(x)| \leq L\|y - x\|$, where $L = \frac{2(M-m)}{R}$. $\square$

Part II

Differentiability of convex functions on $\mathbb{R}^d$

First-Order Regularity

We cannot expect a convex function to be differentiable everywhere.
Example: $y = |x|$ is not differentiable at $x = 0$.
However, convexity prevents this issue from occurring at too many points.

Theorem 2

ϕ is differentiable **almost everywhere (a.e.)** on $\mathbb{R}^d$.
For a.e. $x \in \mathbb{R}^d$, there exists $\nabla\phi(x) \in \mathbb{R}^d$ such that:

$$\phi(y) = \phi(x) + \langle \nabla\phi(x), y - x \rangle + o(\|x - y\|)$$

Remark

Since we proved that ϕ is locally Lipschitz, Rademacher's Theorem immediately implies its differentiability almost everywhere.

First-Order Property

Theorem 2

$\nabla\phi$ is a monotone operator on $\mathbb{R}^d$:
For a.e. $x, y \in \mathbb{R}^d$,

$$\langle \nabla\phi(x) - \nabla\phi(y), x - y \rangle \geq 0$$

Remark

In fact, ϕ is convex if and only if $\nabla\phi$ is define a.e. and monotone.

A Geometric Perspective: Subdifferentials

Property (Supporting Hyperplanes)

Let ϕ a differentiable function define on $\mathbb{R}^d$
ϕ is convex at x if and only if:

$$\forall y \in \mathbb{R}^d \quad \phi(y) \geq \phi(x) + \langle \nabla \phi(x), y - x \rangle$$

(The function lies above its tangent lines).

Definition: The Subdifferential

For all $x \in \mathbb{R}^d$, we define the subdifferential $\partial \phi(x)$ of ϕ at x as:

$$\begin{aligned} \partial \phi(x) &:= \{p \in \mathbb{R}^d \mid \forall y \in \mathbb{R}^d,\ \phi(y) \geq \phi(x) + \langle p, y - x \rangle\} \\ &= \{\text{slopes of hyperplanes lying below Graph}(\phi) \text{ touching at } x\} \end{aligned}$$

Properties of the Subdifferential

Proposition

ϕ is differentiable at x **if and only if** $\partial\phi(x)$ is a singleton.
(In this case, $\partial\phi(x) = \{\nabla\phi(x)\}$).

Examples:

- For $\phi(x) = x^2$ at $x = 0$: $\partial\phi(0) = \{0\}$
- For $\phi(x) = |x|$ at $x = 0$: $\partial\phi(0) = [-1, 1]$

Remark

By geometric Hahn Banach theorem, $\partial\phi(x)$ is never empty.

Proposition (Automatic Continuity of the Gradient)

If $x_n \to x \in \mathbb{R}^d$, and both $\nabla\phi(x_n)$ and $\nabla\phi(x)$ exist, then $\nabla\phi(x_n) \to \nabla\phi(x)$.

Use of the Subdifferential

Proposition (Automatic Continuity of the Gradient)

If $x_n \to x \in \mathbb{R}^d$, and both $\nabla\phi(x_n)$ and $\nabla\phi(x)$ exist, then $\nabla\phi(x_n) \to \nabla\phi(x)$.
So ϕ is continuously differentiable a.e.

Part III

Higher-Order Regularity and Quantitative Estimates

Second-Order Regularity: Alexandrov's Theorem

Can we achieve more regularity? Yes!

Theorem (Alexandrov)

ϕ is **twice differentiable** almost everywhere on $\mathbb{R}^d$ in the following sense:
For a.e. $x \in \mathbb{R}^d$, there exists $D^2\phi(x) \in \mathcal{S}_d^+(\mathbb{R})$ such that:

$$\phi(y) = \phi(x) + \langle \nabla\phi(x), y - x \rangle + \frac{1}{2}\langle D^2\phi(x)(y - x), y - x \rangle + o(\|x - y\|^2)$$

Remarks

- This is a generalized, pointwise second-order Taylor expansion.
- In fact, if a function ϕ is twice differentiable, ϕ is convex iif $D^2\phi(x) \in \mathcal{S}_d^+(\mathbb{R})$ iif ϕ vérifies the last inequality.
- The proof is significantly harder than for first-order differentiability.

Towards Alexandrov: A Quantitative Lemma

I will not present the full proof of Alexandrov's theorem. To give a taste of the underlying ideas, I will present a lemma regarding the Lipschitz regularity of $\nabla\phi$.

Main Lemma: Lipschitz Control of $\nabla\phi$

Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be convex and globally Lipschitz. For all $R > 0$ and $0 < \eta \leq 1$:

$$\sup_{\|h\|<\eta} \int_{B_R} \|\nabla\phi(x+h) - \nabla\phi(x)\| dx \leq C_{d,R} \, \mathrm{Lip}(\phi) \, \eta$$

Remark

This provides a quantitative control on the Lipschitz constant of $\nabla\phi$, but only in the mean (L^1 sense). This is enough to prove Alexandrov's theorem.

A Key Ingredient: Reverse Hölder Inequality

Lemma Bis (Reverse Hölder Inequality)

For any $x \in \mathbb{R}^d$ and $\eta > 0$:

$$\|\nabla \phi\|_{L^\infty(B(x,\eta))} \leq \frac{C_d}{\eta^d} \|\nabla \phi\|_{L^1(B(x,4\eta))}$$

Proof Setup

Assume $\nabla \phi$ exists everywhere (if not, use a smooth convex approximation argument).
By translation invariance, we can assume $x = 0$.
Let $y \in \overline{B}(0, \eta)$ be the point where the supremum is achieved:

$$\|\nabla \phi\|_{L^\infty(B(0,\eta))} = \|\nabla \phi(y)\|$$

Let $e = \frac{\nabla \phi(y)}{\|\nabla \phi(y)\|} \in \mathbb{S}^{d-1}$ be the normalized gradient direction.

Proof of Lemma Bis (Reverse Hölder) - Part 2/3

Exploiting Monotonicity

Since $\nabla\phi$ is monotone, for all $y' \in \mathbb{R}^d$:

$$\langle \nabla\phi(y') - \nabla\phi(y), y' - y \rangle \geq 0$$

Let $B' = B(y + 2\eta e, \eta)$. Notice that $B' \subseteq B(0, 4\eta)$.
Take any $y' \in B'$, which can be written as $y' = y + 2\eta e + z$ with $\|z\| \leq \eta$.
Substituting y' into the monotonicity inequality:

$$\begin{aligned}
\langle \nabla\phi(y'), 2\eta e + z \rangle &\geq \langle \nabla\phi(y), 2\eta e + z \rangle \\
&\geq 2\eta\|\nabla\phi(y)\| - \|\nabla\phi(y)\|\|z\| \\
&\geq 2\eta\|\nabla\phi(y)\| - \eta\|\nabla\phi(y)\| \\
&= \eta\|\nabla\phi(y)\|
\end{aligned}$$

Proof of Lemma Bis (Reverse Hölder) - Part 3/3

Integration Step

Using Cauchy-Schwarz on the left side of the previous inequality:

$$\eta\|\nabla\phi(y)\| \leq \langle \nabla\phi(y'), 2\eta e + z\rangle \leq \|\nabla\phi(y')\|(3\eta)$$

Thus, $\|\nabla\phi(y)\| \leq 3\|\nabla\phi(y')\|$ for all $y' \in B'$.
Integrating this constant over the ball B':

$$\begin{aligned} \mathrm{Vol}(B')\|\nabla\phi(y)\| &\leq \int_{B'} 3\|\nabla\phi(y')\|dy' \\ &\leq 3\|\nabla\phi\|_{L^1(B(0,4\eta))} \end{aligned}$$

Since $\mathrm{Vol}(B') = C_d\eta^d$, the result follows immediately. $\square$

Proof of the Main Lemma - Part 1/4

Applying Lemma Bis

Fix $h \in \mathbb{R}^d$ such that $\|h\| < \eta$. We want to bound:

$$I = \int_{B_R} \|\nabla\phi(x+h) - \nabla\phi(x)\| dx$$

Notice that $\|\nabla\phi(x+h) - \nabla\phi(x)\| \leq \|\nabla\phi(\cdot) - \nabla\phi(x)\|_{L^\infty(B(x,\eta))}$.
We apply Lemma Bis to the shifted convex function $\tilde{\phi}(\cdot) = \phi(\cdot) - \langle\nabla\phi(x), \cdot\rangle$, where $\nabla\tilde{\phi}(\cdot) = \nabla\phi(\cdot) - \nabla\phi(x)$:

$$I \leq \int_{B_R} \frac{C_d}{\eta^d} \|\nabla\phi(\cdot) - \nabla\phi(x)\|_{L^1(B(x,4\eta))} dx$$

Proof of the Main Lemma - Part 2/4

Fubini and the Hessian

Expanding the L^1 norm and swapping the integrals via Fubini's theorem:

$$\begin{aligned} I &\leq \frac{C_d}{\eta^d} \int_{B_{4\eta}} \left(\int_{B_R} \|\nabla\phi(x+u) - \nabla\phi(x)\| dx \right) du \\ &= \frac{C_d}{\eta^d} \int_{B_{4\eta}} \int_{B_R} \left\| \int_0^1 D^2\phi(x+tu)u \, dt \right\| dx \, du \\ &\leq \frac{C_d}{\eta^d} \int_{B_{4\eta}} \|u\| \left(\int_{B_{R+4\eta}} \|D^2\phi(w)\| dw \right) du \end{aligned}$$

Proof of the Main Lemma - Part 3/4

The Laplacian Trick Integration by Parts

Since $D^2\phi \in \mathcal{S}_d^+(\mathbb{R})$, its spectral norm is bounded by its trace (the Laplacian):

$$\|D^2\phi\| = \max_{\lambda \in Sp(D^2\phi)} \lambda \leq \mathrm{tr}(D^2\phi) = \Delta\phi$$

We bound the integral of the Hessian by the integral of the Laplacian using the Divergence Theorem on the extended ball:

$$\begin{aligned} \int_{B_{R+4\eta}} \|D^2\phi\| &\leq \int_{B_{R+4\eta}} \Delta\phi \\ &= \int_{\partial B_{R+4\eta}} \langle \nabla\phi(s), n(s) \rangle d\sigma(s) \\ &\leq \mathrm{Vol}(\partial B_{R+4\eta}) \|\nabla\phi\|_\infty \\ &\leq C_d (R+4\eta)^{d-1} \mathrm{Lip}(\phi) \end{aligned}$$

Proof of the Main Lemma - Part 4/4

Final Integration

We plug the constant $C_{d,R}\text{Lip}(\phi)$ back into our main integral. The only remaining integration is over the vector u:

$$\begin{aligned} I &\leq \frac{C_{d,R}\text{Lip}(\phi)}{\eta^d} \int_{B_{4\eta}} \|u\| du \\ &\leq \frac{C_{d,R}\text{Lip}(\phi)}{\eta^d} \left(C_d \eta^{d+1} \right) \\ &= C'_{d,R}\, \text{Lip}(\phi)\, \eta \end{aligned}$$

Taking the supremum over all $\|h\| < \eta$ gives the desired bound. □

Thanks for your attention !