Gradient Flows for λ-Convex Functions

Kenneth DeMason

University of Texas at Austin

April 17, 2022

Outline

i) Basics of gradient flows of λ-convex functions in $\mathbb{R}^n$.
ii) A priori estimates.
iii) Uniqueness and existence of gradient flows.
iv) Equivalent perspectives of gradient flows.
v) Generalizations to metric spaces.

λ-Convex Functions:

Basically just generalizations of convex functions.

Definition (λ-convexity)

A map $\varphi \in C^2(\mathbb{R}^n)$ is λ-convex if one of the following (equivalent) properties holds

i) $\nabla^2\varphi \geq \lambda\,\mathrm{Id}$ (i.e., for all $\xi \in \mathbb{R}^n$ it holds that $\langle \nabla^2\varphi[\xi], \xi\rangle \geq \lambda|\xi|^2$).

ii) $\langle \nabla\varphi(x_0) - \nabla\varphi(x_1), x_0 - x_1\rangle \geq \lambda|x_0 - x_1|^2$.

iii) $\varphi(x_t) \leq (1-t)\varphi(x_0) + t\varphi(x_1) + \frac{\lambda}{2}t(1-t)|x_1 - x_0|^2$, where $x_t := (1-t)x_0 + tx_1$ and $t \in [0,1]$.

iv) $\varphi(x_1) - \varphi(x_0) \geq \langle \nabla\varphi(x_0), x_1 - x_0\rangle + \frac{\lambda}{2}|x_1 - x_0|^2$.

Remark

φ is λ-convex when $\varphi(x) - \lambda/2|x|^2$ is convex. In particular, φ is convex exactly when $\lambda = 0$.

λ-Convex Functions:

Basically just generalizations of convex functions.

Definition (λ-convexity)

A map $\varphi \in C^2(\mathbb{R}^n)$ is λ-convex if one of the following (equivalent) properties holds

i) $\nabla^2 \varphi \geq \lambda \, \mathrm{Id}$ (i.e., for all $\xi \in \mathbb{R}^n$ it holds that $\langle \nabla^2 \varphi[\xi], \xi \rangle \geq \lambda |\xi|^2$).

ii) $\langle \nabla \varphi(x_0) - \nabla \varphi(x_1), x_0 - x_1 \rangle \geq \lambda |x_0 - x_1|^2$.

iii) $\varphi(x_t) \leq (1-t)\varphi(x_0) + t\varphi(x_1) + \frac{\lambda}{2} t(1-t)|x_1 - x_0|^2$, where $x_t := (1-t)x_0 + tx_1$ and $t \in [0,1]$.

iv) $\varphi(x_1) - \varphi(x_0) \geq \langle \nabla \varphi(x_0), x_1 - x_0 \rangle + \frac{\lambda}{2}|x_1 - x_0|^2$.

Remark

φ is λ-convex when $\varphi(x) - \lambda/2|x|^2$ is convex. In particular, φ is convex exactly when $\lambda = 0$.

λ-Convex Functions:

Basically just generalizations of convex functions.

Definition (λ-convexity)

A map $\varphi \in C^2(\mathbb{R}^n)$ is λ-convex if one of the following (equivalent) properties holds

i) $\nabla^2\varphi \geq \lambda\,\mathrm{Id}$ (i.e., for all $\xi \in \mathbb{R}^n$ it holds that $\langle \nabla^2\varphi[\xi], \xi\rangle \geq \lambda|\xi|^2$).

ii) $\langle \nabla\varphi(x_0) - \nabla\varphi(x_1), x_0 - x_1\rangle \geq \lambda|x_0 - x_1|^2$.

iii) $\varphi(x_t) \leq (1-t)\varphi(x_0) + t\varphi(x_1) + \frac{\lambda}{2}t(1-t)|x_1 - x_0|^2$, where $x_t := (1-t)x_0 + tx_1$ and $t \in [0,1]$.

iv) $\varphi(x_1) - \varphi(x_0) \geq \langle \nabla\varphi(x_0), x_1 - x_0\rangle + \frac{\lambda}{2}|x_1 - x_0|^2$.

Remark

φ is λ-convex when $\varphi(x) - \lambda/2|x|^2$ is convex. In particular, φ is convex exactly when $\lambda = 0$.

Gradient Flows:

Definition (Gradient Flows)

The gradient flow of φ with initial datum $u_0 \in \mathbb{R}^n$ is the unique $C^1((0,\infty))$ solution to

$$\begin{cases} u'(t) = -\nabla\varphi(u(t)) \\ u(0^+) = u_0 \end{cases}$$

Do solutions exist? How do we know the exist for long-time? What about uniqueness?

Gradient Flows:

Definition (Gradient Flows)

The gradient flow of φ with initial datum $u_0 \in \mathbb{R}^n$ is the unique $C^1((0,\infty))$ solution to

$$\begin{cases} u'(t) = -\nabla\varphi(u(t)) \\ u(0^+) = u_0 \end{cases}$$

Do solutions exist? How do we know the exist for long-time? What about uniqueness?

A priori Estimates for Gradient Flows:

Theorem

Given $\varphi \in C^2(\mathbb{R}^n)$ a λ-convex function, and u a gradient flow of φ, we have

EVI: $\frac{1}{2}\frac{d}{dt}|u(t)-v|^2 + \frac{\lambda}{2}|u(t)-v|^2 \leq \varphi(v) - \varphi(u(t))$ *for any* $v \in \mathbb{R}^n$*;*

EI: $\frac{d}{dt}\varphi(u(t)) = -|u'(t)|^2 = -|\nabla\varphi(u(t))|^2 \leq 0$*;*

SI: $\frac{d}{dt}(e^{2\lambda t}|\nabla\varphi(u(t))|^2) = \frac{d}{dt}(e^{2\lambda t}|u'(t)|^2) \leq 0$*.*

Moreover, if v is another gradient flow of φ (with possibly different initial datum) then

Cont $\frac{d}{dt}(e^{\lambda t}|u(t)-v(t)|) \leq 0$*.*

A priori Estimates for Gradient Flows:

Proof.

EVI: Using the subgradient property iv),

$$\frac{1}{2}\frac{d}{dt}|u(t)-v|^2 = \langle u'(t), u(t)-v\rangle = -\langle \nabla\varphi(u(t)), u(t)-v\rangle$$
$$\leq \varphi(v)-\varphi(u(t))-\frac{\lambda}{2}|u(t)-v|^2.$$

Cont: Using the monotonicity of the gradient,

$$\frac{d}{dt}|u(t)-v(t)|^2 = 2\langle u'(t)-v'(t), u(t)-v(t)\rangle$$
$$= -2\langle \nabla\varphi(u(t))-\nabla\varphi(v(t)), u(t)-v(t)\rangle$$
$$\leq -2\lambda|u(t)-v(t)|^2.$$

□

A priori Estimates for Gradient Flows:

With the evolution variational inequality we can derive uniqueness via a "doubling of variables" technique. I.e., for gradient flows $u(t)$ and $v(s)$ we have

$$\frac{1}{2}\frac{d}{dt}|u(t) - v(s)|^2 + \frac{\lambda}{2}|u(t) - v(s)|^2 \leq \varphi(v(s)) - \varphi(u(t))$$
$$\frac{1}{2}\frac{d}{ds}|v(s) - u(t)|^2 + \frac{\lambda}{2}|u(t) - v(s)|^2 \leq \varphi(u(t)) - \varphi(v(s))$$

Adding these and setting $t = s$ implies

$$\frac{d}{dt}|u(t) - v(t)|^2 \leq 0,$$

so if $u(t)$ and $v(t)$ start at the same point, they must remain the same.

A priori Estimates for Gradient Flows:

With the evolution variational inequality we can derive uniqueness via a "doubling of variables" technique. I.e., for gradient flows $u(t)$ and $v(s)$ we have

$$\frac{1}{2}\frac{d}{dt}|u(t)-v(s)|^2+\frac{\lambda}{2}|u(t)-v(s)|^2\leq\varphi(v(s))-\varphi(u(t))$$
$$\frac{1}{2}\frac{d}{ds}|v(s)-u(t)|^2+\frac{\lambda}{2}|u(t)-v(s)|^2\leq\varphi(u(t))-\varphi(v(s))$$

Adding these and setting $t=s$ implies

$$\frac{d}{dt}|u(t)-v(t)|^2\leq 0,$$

so if $u(t)$ and $v(t)$ start at the same point, they must remain the same.

A priori Estimates for Gradient Flows:

Corollary

If u and v are gradient flows for the λ-convex function φ, then

$$|u(t) - v(t)| \leq e^{-\lambda t}|u_0 - v_0|$$

Proof.

Gronwall's lemma applied to the contraction property. □

Importantly, this shows uniqueness of gradient flows with the same initial datum.

A priori Estimates for Gradient Flows:

Corollary

If u and v are gradient flows for the λ-convex function φ, then

$$|u(t) - v(t)| \leq e^{-\lambda t}|u_0 - v_0|$$

Proof.

Gronwall's lemma applied to the contraction property. □

Importantly, this shows uniqueness of gradient flows with the same initial datum.

A Semigroup Perspective:

For each initial datum u_0 we define $S_t(u_0)$ to be the unique gradient flow of φ. Then, $\{S_t\}_{t>0}$ is a contractive semigroup. That is,

$$S_t S_h(u_0) = S_{t+h}(u_0)$$

with $S_t(u_0) \to u_0$ as $t \to 0^+$. By the previous contraction property,

$$|S_t(u_0) - S_t(v_0)| \leq e^{-\lambda t}|u_0 - v_0|.$$

This shows, for example, uniqueness and continuous dependence with respect to the initial datum.

A Semigroup Perspective:

For each initial datum u_0 we define $S_t(u_0)$ to be the unique gradient flow of φ. Then, $\{S_t\}_{t>0}$ is a contractive semigroup. That is,

$$S_t S_h(u_0) = S_{t+h}(u_0)$$

with $S_t(u_0) \to u_0$ as $t \to 0^+$. By the previous contraction property,

$$|S_t(u_0) - S_t(v_0)| \leq e^{-\lambda t}|u_0 - v_0|.$$

This shows, for example, uniqueness and continuous dependence with respect to the initial datum.

A Semigroup Perspective:

For each initial datum u_0 we define $S_t(u_0)$ to be the unique gradient flow of φ. Then, $\{S_t\}_{t>0}$ is a contractive semigroup. That is,

$$S_t S_h(u_0) = S_{t+h}(u_0)$$

with $S_t(u_0) \to u_0$ as $t \to 0^+$. By the previous contraction property,

$$|S_t(u_0) - S_t(v_0)| \leq e^{-\lambda t}|u_0 - v_0|.$$

This shows, for example, uniqueness and continuous dependence with respect to the initial datum.

Approximating Solutions:

To show existence, we construct approximate solutions and show they converge to a C^1 solutions. The approximation scheme is known as backwards (or implicit) Euler.

Let $\tau > 0$ and partition $[0, \infty)$ as

$$\mathcal{P}_\tau = \{0 = t_\tau^0 < t_\tau^1 < .. < t_\tau^n < ...\}$$

with $t_\tau^n = n\tau$.

We search for a sequence $\{U_\tau^n\}_{n=0}^\infty$ such that $U_\tau^n \approx u(t_\tau^n)$. By defining $\bar{U}_\tau(t) = U_\tau^n$ on $(t_\tau^{n-1}, t_\tau^n]$, we hope that $\bar{U}_\tau(t)$ converges to a gradient flow u.

Approximating Solutions:

To show existence, we construct approximate solutions and show they converge to a C^1 solutions. The approximation scheme is known as backwards (or implicit) Euler.

Let $\tau > 0$ and partition $[0, \infty)$ as

$$\mathcal{P}_\tau = \{0 = t_\tau^0 < t_\tau^1 < .. < t_\tau^n < ...\}$$

with $t_\tau^n = n\tau$.

We search for a sequence $\{U_\tau^n\}_{n=0}^\infty$ such that $U_\tau^n \approx u(t_\tau^n)$. By defining $\bar{U}_\tau(t) = U_\tau^n$ on $(t_\tau^{n-1}, t_\tau^n]$, we hope that $\bar{U}_\tau(t)$ converges to a gradient flow u.

Approximating Solutions:

To show existence, we construct approximate solutions and show they converge to a C^1 solutions. The approximation scheme is known as backwards (or implicit) Euler.

Let $\tau > 0$ and partition $[0, \infty)$ as

$$\mathcal{P}_\tau = \{0 = t_\tau^0 < t_\tau^1 < .. < t_\tau^n < ...\}$$

with $t_\tau^n = n\tau$.

We search for a sequence $\{U_\tau^n\}_{n=0}^\infty$ such that $U_\tau^n \approx u(t_\tau^n)$. By defining $\bar{U}_\tau(t) = U_\tau^n$ on $(t_\tau^{n-1}, t_\tau^n]$, we hope that $\bar{U}_\tau(t)$ converges to a gradient flow u.

Approximating Solutions:

We construct U_τ^n iteratively, starting with $U_\tau^0 = u_0$. Given U_τ^{n-1} define U_τ^n by solving

$$\frac{U_\tau^n - U_\tau^{n-1}}{\tau} = -\nabla\varphi(U_\tau^n).$$

Solutions exist by looking for minimizers of the functional

$$v \mapsto \Phi(\tau, U_\tau^{n-1}; v) := \frac{1}{2\tau}|v - U_\tau^{n-1}|^2 + \varphi(v).$$

Since Φ is $(\tau^{-1} + \lambda)$-convex, it admits a unique minimizer whenever $\tau^{-1} > -\lambda$. In particular, if $\lambda \geq 0$ we always have unique minimizers.

Approximating Solutions:

We construct U_τ^n iteratively, starting with $U_\tau^0 = u_0$. Given U_τ^{n-1} define U_τ^n by solving

$$\frac{U_\tau^n - U_\tau^{n-1}}{\tau} = -\nabla\varphi(U_\tau^n).$$

Solutions exist by looking for minimizers of the functional

$$v \mapsto \Phi(\tau, U_\tau^{n-1}; v) := \frac{1}{2\tau}|v - U_\tau^{n-1}|^2 + \varphi(v).$$

Since Φ is $(\tau^{-1} + \lambda)$-convex, it admits a unique minimizer whenever $\tau^{-1} > -\lambda$. In particular, if $\lambda \geq 0$ we always have unique minimizers.

Approximating Solutions:

We construct U_τ^n iteratively, starting with $U_\tau^0 = u_0$. Given U_τ^{n-1} define U_τ^n by solving

$$\frac{U_\tau^n - U_\tau^{n-1}}{\tau} = -\nabla\varphi(U_\tau^n).$$

Solutions exist by looking for minimizers of the functional

$$v \mapsto \Phi(\tau, U_\tau^{n-1}; v) := \frac{1}{2\tau}|v - U_\tau^{n-1}|^2 + \varphi(v).$$

Since Φ is $(\tau^{-1} + \lambda)$-convex, it admits a unique minimizer whenever $\tau^{-1} > -\lambda$. In particular, if $\lambda \geq 0$ we always have unique minimizers.

The Generalized Setting:

Before moving into a more general framework, we'll discuss two approaches to viewing the gradient flow.

- Curves of maximal slope,
- Evolution variational inequalities.

Both of these ideas rest on characterizing the gradient flow equation in a different way, which can be adapted to a metric setting.

The Generalized Setting:

Before moving into a more general framework, we'll discuss two approaches to viewing the gradient flow.

- Curves of maximal slope,
- Evolution variational inequalities.

Both of these ideas rest on characterizing the gradient flow equation in a different way, which can be adapted to a metric setting.

Curves of Maximal Slope:

Based off of the energy inequality.

Proposition (Curves of maximal slope)

A C^1 curve $u : [0, \infty) \to \mathbb{R}^n$ is a gradient flow of φ if and only if it satisfies the energy dissipation inequality

$$\frac{d}{dt}\varphi(u(t)) \leq -\frac{1}{2}|u'(t)|^2 - \frac{1}{2}|\nabla\varphi(u(t))|^2$$

or the weaker integrated form

$$\varphi(u(t)) + \frac{1}{2}\int_0^t \left(|u'(s)|^2 + |\nabla\varphi(s)|^2\right) ds \leq \varphi(u_0).$$

Curves of Maximal Slope:

Proof.

If u is a C^1 curve, then by the chain rule

$$\varphi(u(t)) = \varphi(u_0) + \int_0^t \langle \nabla\varphi(u(s)), u'(s)\rangle \, ds.$$

By applying the (weak) energy dissipation inequality, we have

$$\frac{1}{2}\int_0^t |u'(s) + \nabla\varphi(u(s))|^2 \, ds = \frac{1}{2}\int_0^t \left(|u'(s)|^2 + |\nabla\varphi(u(s))|^2\right) ds + \int_0^t \langle \nabla\varphi(u(s)), u'(s)\rangle \, ds \leq 0.$$

That is, $u'(s) = -\nabla\varphi(u(s))$ for $\mathcal{L}^1$-a.e. $s \in (0, t)$, for all $t > 0$. □

Note: We did not use λ-convexity!

Curves of Maximal Slope:

Proof.

If u is a C^1 curve, then by the chain rule

$$\varphi(u(t)) = \varphi(u_0) + \int_0^t \langle \nabla\varphi(u(s)), u'(s)\rangle \, ds.$$

By applying the (weak) energy dissipation inequality, we have

$$\begin{aligned}\frac{1}{2}\int_0^t |u'(s) + \nabla\varphi(u(s))|^2 \, ds = \frac{1}{2}\int_0^t \left(|u'(s)|^2 + |\nabla\varphi(u(s))|^2\right) ds \\ + \int_0^t \langle \nabla\varphi(u(s)), u'(s)\rangle \, ds \leq 0.\end{aligned}$$

That is, $u'(s) = -\nabla\varphi(u(s))$ for $\mathcal{L}^1$-a.e. $s \in (0, t)$, for all $t > 0$. □

Note: We did not use λ-convexity!

Curves of Maximal Slope:

Proof.

If u is a C^1 curve, then by the chain rule

$$\varphi(u(t)) = \varphi(u_0) + \int_0^t \langle \nabla\varphi(u(s)), u'(s)\rangle \; ds.$$

By applying the (weak) energy dissipation inequality, we have

$$\frac{1}{2}\int_0^t |u'(s) + \nabla\varphi(u(s))|^2 \; ds = \frac{1}{2}\int_0^t \left(|u'(s)|^2 + |\nabla\varphi(u(s))|^2\right) ds + \int_0^t \langle \nabla\varphi(u(s)), u'(s)\rangle \; ds \leq 0.$$

That is, $u'(s) = -\nabla\varphi(u(s))$ for $\mathcal{L}^1$-a.e. $s \in (0, t)$, for all $t > 0$. □

Note: We did not use λ-convexity!

Evolution Variational Inequality:

Similar to the previous, we have

Proposition (Curves of maximal slope)

A C^1 curve $u : [0, \infty) \to \mathbb{R}^n$ satisfying the EVI is a gradient flow of φ.

Proof.

By differentiating the norm in the EVI,

$$\langle u'(t), u(t) - v \rangle \leq \varphi(v) - \varphi(u(t)) - \frac{\lambda}{2}|v - u(t)|^2.$$

For $\xi \in \mathbb{R}^n$ and $\epsilon > 0$ set $v = u(t) + \epsilon\xi$. Then the above reads

$$-\epsilon\langle u'(t), \xi \rangle = \langle \nabla\varphi(u(t)), v - u(t) \rangle \leq \varphi(u(t) + \epsilon\xi) - \varphi(u(t)) - \frac{\lambda\epsilon^2}{2}|\xi|^2.$$

Dividing by ϵ and taking $\epsilon \to 0$ concludes. □

Approaches in Generalized Settings:

There are four main approaches we will (briefly) discuss.

i) Curves of maximal slope in metric spaces.
ii) Generalized minimizing movements
iii) Differential inclusions and Hilbert spaces.
iv) Evolution variational inequalities in metric spaces.

For all, we need to relax our notion of derivative (since in a metric setting we do not generally have access to a differential structure). This is known as metric calculus.

Approaches in Generalized Settings:

There are four main approaches we will (briefly) discuss.

i) Curves of maximal slope in metric spaces.

ii) Generalized minimizing movements

iii) Differential inclusions and Hilbert spaces.

iv) Evolution variational inequalities in metric spaces.

For all, we need to relax our notion of derivative (since in a metric setting we do not generally have access to a differential structure). This is known as metric calculus.

The Basics of Metric Calculus

Typically we work with a complete, separable metric space (X, d) and $\phi : (-\infty, \infty]$ a proper, lower-semicontinuous functional. The idea of this theory is to work with curves. For example, if ϕ is differentiable then

$$|\nabla \phi| \leq g \quad \Leftrightarrow \quad |(\phi \circ v)'| \leq g(v)|v'|$$

for any regular curve v on $\mathbb{R}^n$.

Definition (Metric derivative)

Given an absolutely continuous curve $v : (a, b) \to X$, the metric derivative is defined by

$$|v'|(t) = \lim_{s \to t} \frac{d(v(t), v(s))}{|t - s|}.$$

The Basics of Metric Calculus

Typically we work with a complete, separable metric space (X, d) and $\phi : (-\infty, \infty]$ a proper, lower-semicontinuous functional. The idea of this theory is to work with curves. For example, if ϕ is differentiable then

$$|\nabla \phi| \leq g \quad \Leftrightarrow \quad |(\phi \circ v)'| \leq g(v)|v'|$$

for any regular curve v on $\mathbb{R}^n$.

Definition (Metric derivative)

Given an absolutely continuous curve $v : (a, b) \to X$, the metric derivative is defined by

$$|v'|(t) = \lim_{s \to t} \frac{d(v(t), v(s))}{|t - s|}.$$

The Basics of Metric Calculus

Definition (Metric Slope)

The metric slope of ϕ at a point $v \in X$ is defined by

$$|\partial\phi|(v) = \limsup_{w \to v} \frac{(\phi(v) - \phi(w))^+}{d(v, w)}.$$

The metric slope is an example of an upper gradient g defined previously.

These allow one to replace derivatives in the energy dissipation inequality and evolution variational inequality with the metric derivatives and slope. By finding curves which satisfy these inequalities, we say that the curve is a gradient flow (using the equivalent formulation in the smooth case as justification).

The Basics of Metric Calculus

Definition (Metric Slope)

The metric slope of ϕ at a point $v \in X$ is defined by

$$|\partial\phi|(v) = \limsup_{w \to v} \frac{(\phi(v) - \phi(w))^+}{d(v, w)}.$$

The metric slope is an example of an upper gradient g defined previously.

These allow one to replace derivatives in the energy dissipation inequality and evolution variational inequality with the metric derivatives and slope. By finding curves which satisfy these inequalities, we say that the curve is a gradient flow (using the equivalent formulation in the smooth case as justification).