* Exponential stability: $\dot V \le -\alpha V$, $\alpha > 0$

$$\alpha_1 \|x\|^2 \le V \le \alpha_2 \|x\|^2 \qquad V(0)=0$$
$$\dot V \le -\alpha_3 \|x\|^2$$

$$m \le \sqrt{\frac{\alpha_2}{\alpha_1}}, \quad \alpha \ge \frac{\alpha_3}{2\alpha_2}$$

$$\|x(t)\| \le m e^{-\alpha(t-t_0)}$$

* Assume: f is μ-SC | $\nabla^2 f \succeq \mu I$

$$\boxed{\dot x = -\nabla f(x)}$$
$\hookrightarrow$ GF

$$V = \frac{1}{2}\|\nabla f\|^2$$

only at $x = x^*$, $\nabla f(x^*) = 0$

$V(x^*) = 0$

$V(x) > 0 \quad \forall x \setminus \{x^*\}$

$$\dot V = (\nabla f)^T \nabla^2 f \, \dot x$$
$$\dot V = -(\nabla f)^T \nabla^2 f \, (\nabla f)$$
$$\le -\mu \|\nabla f\|^2$$
$$\dot V \le -2\mu V$$

$\Longrightarrow$ V converges exponentially fast

What about x?

GF: • $\dot x = -\nabla f(x)$ with x^* being the optimizer

$\nabla f(x^*) = 0$

• Assume that f is μ-SC and L-smooth.

Since f is μ-SC:

$$f(y) \ge f(x) + \nabla f(x)^T (y - x) + \frac{\mu}{2}\|y - x\|^2$$

Choose $y = x$ and $x = x^*$

$$\boxed{f(x) \ge f^* + \frac{\mu}{2}\|x - x^*\|^2} \quad \text{— ①}$$

f is L-smooth:

$$f(y) \le f(x) + \nabla f(x)^T (y - x) + \frac{L}{2}\|y - x\|^2$$

$$\boxed{f(x) \le f(x^*) + \frac{L}{2}\|x - x^*\|^2} \quad \text{— ②}$$

Combining (1) and (2):

$$\boxed{\frac{\mu}{2}\|x-x^*\|^2 \le \underbrace{f(x)-f^*}_{V} \le \frac{L}{2}\|x-x^*\|^2} \quad \text{— (a)}$$

$$V = f(x) - f^*$$

$$\dot{V} = \nabla f(x)^T \dot{x}$$

$$= -\|\nabla f\|^2 = -2\mu \frac{1}{2\mu}\|\nabla f\|^2 \le -2\mu\underbrace{(f(x)-f^*)}_{V}$$

(PL-inequality)

$$\dot{V} \le -2\mu V \le -2\mu\left(\frac{\mu}{2}\right)\|x-x^*\|^2$$

$$\boxed{\dot{V} \le -\mu^2\|x-x^*\|^2} \quad \text{— (b)}$$

$$\alpha_1 = \frac{\mu}{2}; \quad \alpha_2 = \frac{L}{2}; \quad \alpha_3 = \mu^2$$

$$m \le \sqrt{\frac{\alpha_2}{\alpha_1}} = \sqrt{\frac{L}{\mu}} \qquad \alpha \ge \frac{\alpha_3}{2\alpha_2} = \frac{\mu^2}{L}$$

From (a) and (b):

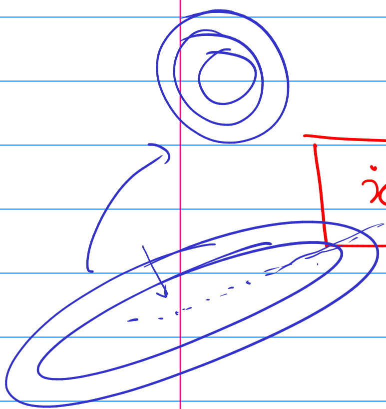
$$\boxed{\|x(t)-x^*\| \le \sqrt{\frac{L}{\mu}}\, e^{-\mu^2/L\, t}}$$

— x — x — x — x — x — x —

* Assume f is convex and $\nabla^2 f > 0 \Rightarrow f$ is strictly convex.

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$$

$$\forall\ x \ne y \quad \lambda \in (0,1)$$



$$\boxed{\dot{x} = -(\nabla^2 f(x))^{-1}\nabla f(x)} \leftarrow \text{Newton's Method}$$

$$V = \frac{1}{2}\|\nabla f\|^2$$

$$\dot{V} = (\nabla f)^T(\nabla^2 f)\,\dot{x}$$

$$= -(\nabla f)^T \nabla^2 f \left\{(\nabla^2 f)^{-1}\nabla f\right\}$$

$$= -\|\nabla f\|^2 = -2V \Rightarrow \dot{V} = -2V$$

$\therefore$ V converges exponentially fast.
However, we cannot guarantee exponential stability of x^*.

—x—x—x—x—x—x—

* Consider the case when f is simply convex.

$\nabla^2 f \geq 0$ (2^{nd} order condition for convexity)

Since $\nabla^2 f$ is not invertible, we use gradient flows.

$$\dot{x} = -\nabla f(x)$$

$$V = \frac{1}{2}\|\nabla f\|^2$$

$$\dot{V} = \nabla f^T \nabla^2 f \, \dot{x} = -(\nabla f)^T \nabla^2 f (\nabla f) \leq 0$$

$\dot{V} \leq 0 \Rightarrow$ Stability

—x—x—x—x—x—

eg: $f(x,y) = \frac{1}{2}x^2 + 0.00005\, y^2$

Initial cond^n (x_0, y_0)
$(10, 5)$

$$\nabla f = \begin{bmatrix} x \\ .0001\, y \end{bmatrix}$$

$\rightarrow$ $x_{k+1} = x_k - \eta \, \nabla_x f(x,y)$

$\rightarrow$ $y_{k+1} = y_k - \eta \, \nabla_y f(x,y)$

$$x_{k+1} = x_k - \eta \frac{\nabla_x f(x,y)}{|\nabla_x f(x,y)| + \varepsilon}$$

$$y_{k+1} = y_k - \eta \frac{\nabla_y f(x,y)}{|\nabla_y f(x,y)| + \varepsilon}$$

Gradient normalization

* Choices of Lyapunov functions:

$$V = f(x) - f^* \qquad V = \frac{1}{2}\|x - x^*\|^2$$

$$V = \frac{1}{2}\|\nabla f(x)\|^2$$

* Bregman divergence: (2016 PNAS by Michael Jordan)

$$D_h(p,q) := \underbrace{h(p) - h(q) - \nabla h(q)^T (p-q)}_{>0 \text{ if } p \neq q}$$

$\to$ h is strictly convex

Rescaled Gradient Flows (R-GF):

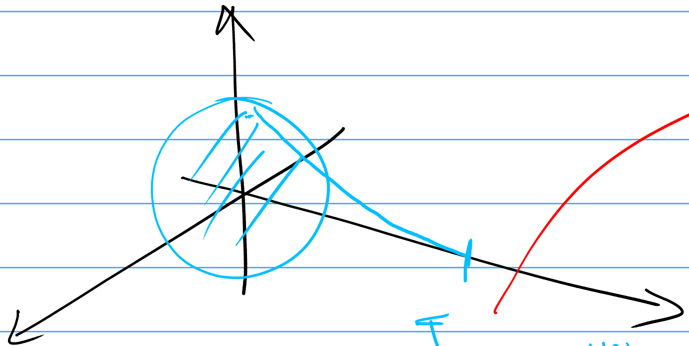
$$\dot{x} = -\frac{\nabla f(x)}{\|\nabla f(x)\|^{\frac{p-2}{p-1}}} \quad , \; p > 2$$

$\to$ RGF

Convergence rate is $O\left(\frac{1}{t^p}\right)$.

Finite-time stability (Bhat 1998)

Fixed-time stability (Polyakov 2012)



T_{x_0} is independent of x_0. (T)

$T < \infty$

$\forall t \geq T$

$x(t) = 0 \;\; \forall x_0$

$T_{x_0} \to$ settling-time function

$T_{x_0} < \infty \quad \forall t \geq T_{x_0}$

$x(t) = 0$

* Lyapunov characterization of finite-time stability:

$\dot{V} \leq -V$ (Exponential convergence of V)

$\dot{V} \leq -V^\alpha \quad 0 < \alpha < 1$ $\to$ Finite-time stability

$$\int_{V_0}^{V(t)} \frac{dV}{V^\alpha} \leq -\int_0^t dt$$

$$\frac{1}{1-\alpha}\left[V(t)^{1-\alpha} - V_0^{1-\alpha}\right] \leq -t$$

$$\underbrace{V(t)^{1-\alpha}}_{\geq 0} \leq V_0^{1-\alpha} - (1-\alpha)t$$

$\exists$ some time $T_{x_0} < \infty$ s.t. $V_0^{1-\alpha} = (1-\alpha) T_{x_0}$

$$\underline{T_{x_0}} = \frac{V_0^{1-\alpha}}{1-\alpha}$$

$$\forall\, t \geq T_{x_0} \quad V(t) = 0$$

Settling-time function, $\boxed{T_{x_0} \leq \dfrac{V(x_0)^{1-\alpha}}{1-\alpha}}$

* Analyzing <u>RGF using Lyapunov Theory</u>:

Assume f is $\underline{\mu\text{-SC}}$:

$$V = \frac{1}{2}\|\nabla f\|^2$$

$$\dot{V} = (\nabla f)^T \nabla^2 f \,\dot{x}$$

$$= -\frac{(\nabla f)^T \nabla^2 f (\nabla f)}{\|\nabla f\|^{\frac{p-2}{p-1}}} \quad , p > 2 \quad (\text{From RGF})$$

$$\leq -\frac{\|\nabla f\|^2 \mu}{\|\nabla f\|^{\frac{p-2}{p-1}}} \quad [\because \nabla^2 f \succeq \mu I]$$

$$= -\mu \|\nabla f\|^{2 - \frac{p-2}{p-1}}$$

$$= -\mu \|\nabla f\|^{\frac{p}{p-1}} = -\mu \|\nabla f\|^{2 \cdot \frac{p}{2(p-1)}}$$

$$= -\mu (2V)^{\frac{p}{2(p-1)}}$$

For $p > 2$, $0 < \dfrac{p}{2(p-1)} < 1$

$$\dot{V} \leq -\mu 2^{\frac{p}{2(p-1)}} V^{\frac{p}{2(p-1)}} \quad ; \alpha = \frac{p}{2(p-1)}$$

$\Rightarrow x^*$ is finite-time stable,

$$T_{x_0} = \frac{V(x_0)^{1-\alpha}}{\mu 2^{\alpha}(1-\alpha)}$$

$\ast$ Fixed-time stability:

$$\boxed{\dot{V} \le -c_1 V^{\alpha_1} - c_2 V^{\alpha_2}}$$

Equilibrium is fixed-time stable

$c_1 > 0$, $\alpha_1 \in (0,1)$
$c_2 > 0$, $\alpha_2 > 1$

Settling-time

$$\boxed{T \le \frac{1}{c_1(1-\alpha_1)} + \frac{1}{c_2(\alpha_2 - 1)}}$$

$\ast$ Another GF (similar to RGF):

$$\dot{x} = -\frac{\nabla f}{\|\nabla f\|^{\frac{p-2}{p-1}}} - \frac{\nabla f}{\|\nabla f\|^{\frac{q-2}{q-1}}}$$

$p > 2$ and $q \in (1,2)$

$$= -\nabla f \left(\frac{1}{\|\nabla f\|^{\frac{p-2}{p-1}}} + \frac{1}{\|\nabla f\|^{\frac{q-2}{q-1}}} \right)$$

$$V = \frac{1}{2} \|\nabla f\|^2$$

$$\dot{V} = (\nabla f)^T \nabla^2 f \, \dot{x}$$

$$= -\frac{(\nabla f)^T \nabla^2 f \, \nabla f}{\|\nabla f\|^{\frac{p-2}{p-1}}} - \frac{(\nabla f)^T \nabla^2 f \, \nabla f}{\|\nabla f\|^{\frac{q-2}{q-1}}}$$

$$\le -\mu (2V)^{\frac{p}{2(p-1)}} - \mu (2V)^{\frac{q}{2(q-1)}}$$

when $p > 2$, $\frac{p}{2(p-1)} \in (0,1)$

$q \in (1,2)$, $\frac{q}{2(q-1)} > 1$

$$\dot{V} \le -c_1 V^{\alpha_1} - c_2 V^{\alpha_2} \qquad T \le \frac{1}{c_1(1-\alpha_1)} + \frac{1}{c_2(\alpha_2 - 1)}$$