Control Tools for Distributed Optimization

The gradient method analysis via control tools

Prof. Ivano Notarnicola

Dept. of Electrical, Electronic, Information Eng.
Università di Bologna
ivano.notarnicola@unibo.it

Prof. Ruggero Carli

Dept. of Information Engineering
Università di Padova
ruggero.carli@unipd.it

SIDRA Ph.D. Summer School
July, 10-12 2025 • Bertinoro, Italy

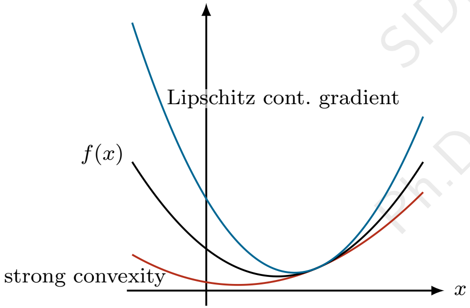
Convex optimization problem (recall)

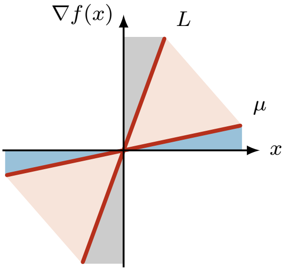
Consider the unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} \; f(x)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is μ-strongly convex and has L-Lipschitz continuous gradient ($L > \mu > 0$)

For all x_A, x_B, it holds $(\nabla f(x_A) - \nabla f(x_B))^\top (x_A - x_B) \geq \frac{1}{\mu+L} \|\nabla f(x_A) - \nabla f(x_B)\|^2 + \frac{\mu L}{\mu+L} \|x_A - x_B\|^2 \geq 0$





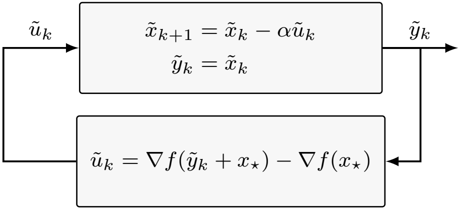
The gradient method in error coordinates

Let $\tilde{x} = x - x_\star$, then the gradient method reads

$$\begin{aligned} \tilde{x}_{k+1} &= \tilde{x}_k - \alpha \tilde{u}_k \\ \tilde{y}_k &= \tilde{x}_k \\ \tilde{u}_k &= \nabla f(\tilde{y}_k + x_\star) - \underbrace{\nabla f(x_\star)} \end{aligned}$$

with $\alpha > 0$ being the so-called *stepsize* and an arbitrary initial condition $x_0 \in \mathbb{R}^n$

The *convergence* analysis amounts to studying the *stability* properties of the origin $\tilde{x} = 0_n$ (unique equilibrium)



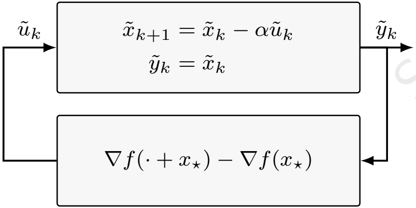
Equivalent representation of the gradient method analysis

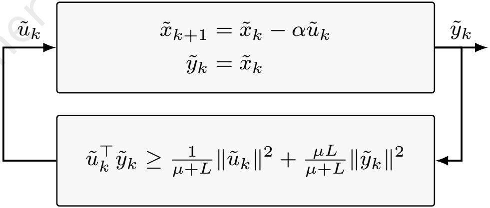
The gradient method is a feedback interconnection of a discrete-time integrator and a static nonlinear map

$$\tilde{y} \longmapsto \tilde{u} = \nabla f(\tilde{y} + x_\star) - \nabla f(x_\star)$$

which can be "replaced" with its sector-bounded characterization ($x_A = \tilde{y}$ and $x_B = 0$)

$$(\nabla f(\tilde{y}_k + x_\star) - \nabla f(x_\star))^\top \tilde{y}_k \geq \frac{1}{\mu+L} \|\nabla f(\tilde{y}_k + x_\star) - \nabla f(x_\star)\|^2 + \frac{\mu L}{\mu+L} \|\tilde{y}_k\|^2$$





Are the two individual components passive?

Gradient method analysis: first loop transformation

First, actuate a (positive) *feedback action* on the plant

$$\tilde{u} = \widehat{u} + \mu \tilde{y}$$

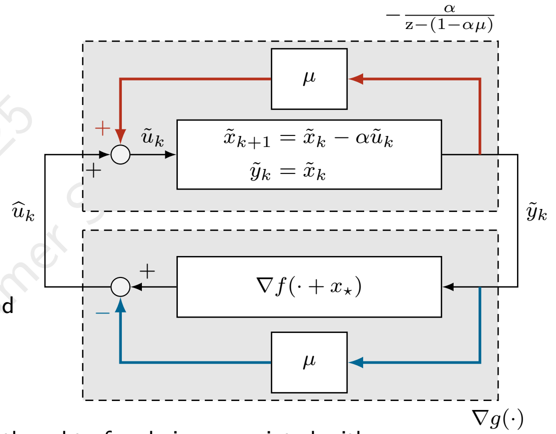
"stealing" strong convexity from the cost function f

The resulting (linear) dynamics is

$$\begin{aligned}\tilde{x}_{k+1} &= (1-\alpha\mu)\tilde{x}_k - \alpha \widehat{u}_k \\ \tilde{y}_k &= \tilde{x}_k\end{aligned}$$

with $\widehat{u}_k = \tilde{u}_k - \mu \tilde{y}_k$ and $\tilde{y}_k$ satisfying the co-coercivity bound

$$\widehat{u}_k^\top \tilde{y}_k \geq \frac{1}{L-\mu} \|\widehat{u}_k\|^2$$



Remark. The static nonlinearity (lower subsystem) can be thought of as being associated with a convex function $g(\cdot)$ having a $(L-\mu)$-Lipschitz continuous gradient

Remark. The transfer function from $\widehat{u}_k$ to $\tilde{y}_k$ is $(1-\mu G_{\tilde{u}\to\tilde{y}}(z))^{-1} G_{\tilde{u}\to\tilde{y}}(z)$, with $G_{\tilde{u}\to\tilde{y}}(z) = -\frac{\alpha}{z-1}$

Gradient method analysis: second loop transformation

Second, consider a *feedforward action* on the plant

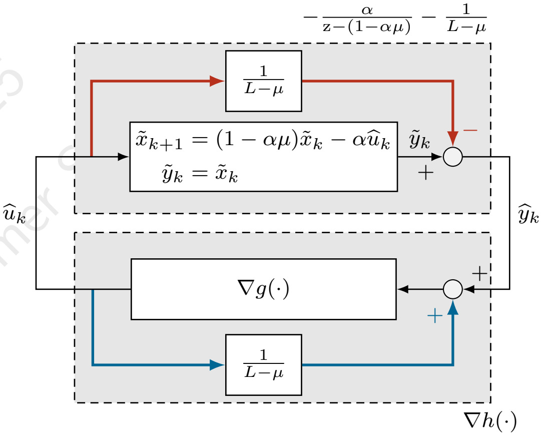
$$\tilde{y} \longmapsto \widehat{y} = \tilde{y} - \frac{1}{L-\mu}\widehat{u}$$

The resulting (linear) dynamics is

$$\begin{aligned}\tilde{x}_{k+1} &= (1-\alpha\mu)\tilde{x}_k - \alpha\widehat{u}_k \\ \widehat{y}_k &= \tilde{x}_k - \frac{1}{L-\mu}\widehat{u}_k\end{aligned}$$

with $\widehat{u}_k$ and $\widehat{y}_k$ satisfying the monotonicity bound

$$\widehat{u}_k^\top \widehat{y}_k \geq 0$$



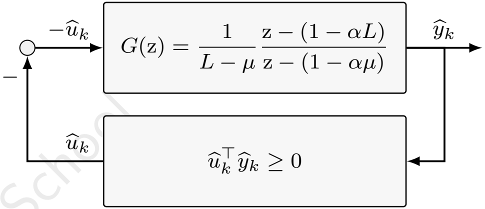
Remark. The static nonlinearity can be thought as associated to a merely convex function $\nabla h(\cdot)$

Gradient method analysis: zero-pole map

The transfer function from $-\widehat{u}_k$ to $\widehat{y}_k$ is

$$G(\mathrm{z}) = \frac{\widehat{Y}(\mathrm{z})}{-\widehat{U}(\mathrm{z})} = \frac{\alpha}{\mathrm{z} - (1-\alpha\mu)} + \frac{1}{L-\mu}$$

$$= \frac{1}{L-\mu}\frac{\mathrm{z}-(1-\alpha L)}{\mathrm{z}-(1-\alpha\mu)}$$



The open-loop plant has a zero at $\mathrm{z} = 1 - \alpha L$ and a pole at $\mathrm{z} = 1 - \alpha\mu$

Recall that $G(\mathrm{z}) + G(\bar{\mathrm{z}})$ is positive outside the disk

- of radius $\frac{|z_0 - p_0|}{2} = \frac{|1-\alpha L - 1 + \alpha\mu|}{2} = \alpha\frac{L-\mu}{2}$ and
- centered at $\mathrm{z} = \frac{z_0 + p_0}{2} = 1 - \alpha\frac{L+\mu}{2}$

For a sufficiently small stepsize $\alpha > 0$, the mentioned disk is entirely contained in the unit disk

Hence $G(\mathrm{z})$ is *strictly* discrete positive real

Remark. For $\alpha > \frac{2}{L}$ the system is non-minimum phase (a real zero less than -1), hence not DPR

Gradient method analysis: convergence proof

Consider a *minimal state-space realization* of the transfer function $G(z)$ given by

$$\tilde{x}_{k+1} = (1-\alpha\mu)\tilde{x}_k + \alpha(-\widehat{u}_k)$$
$$\widehat{y}_k = \tilde{x}_k + \frac{1}{L-\mu}(-\widehat{u}_k)$$

Then, by the KYP lemma, there exists a storage function $V(\tilde{x}) = \frac{1}{2}\tilde{x}^\top P\tilde{x}$, with $P = P^\top > 0$, such that

$$V(\tilde{x}_{k+1}) - V(\tilde{x}_k) = \widehat{y}_k^\top(-\widehat{u}_k) - \frac{1}{2}\|M_y\widehat{y}_k + M_u(-\widehat{u}_k)\|^2 - \big(1-\tfrac{1}{\rho}\big)V(\tilde{x}_k)$$

with $\rho > 1$

When feedback interconnected with a passive nonlinear map (i.e., satisfying $\widehat{y}_k^\top\widehat{u}_k \geq 0$) it follows

$$V(\tilde{x}_{k+1}) - V(\tilde{x}_k) \leq -\big(1-\tfrac{1}{\rho}\big)V(\tilde{x}_k)$$

certifying *exponential stability* of the equilibrium $\tilde{x}=0$, and hence convergence of y_k to the optimal solution $x_\star$

Remark. The explicit values for P, M_y, and M_u are not required