Workshop on Nonsmooth Optimization and Applications NOPTA 8-12 April 2024 University of Antwerp, Belgium

Krasnoselskii-Mann iterations: Inertia, perturbations and approximation

Daniel Cortild, **Juan Peypouquet**

Bernouilli Institute, University of Groningen

d.cortild@rug.nl, j.g.peypouquet@rug.nl

university of groningen

Abstract

Krasnoselskii-Mann iterations constitute a class of fixed point iterations combined with relaxations, employed to approximate fixed points of (quasi-)nonexpansive operators. We present a study of a family of such iterations combining different inertial principles into a single framework. We provide a systematic, unified and insightful analysis of the hypotheses that ensure their weak, strong and linear convergence, either matching or improving previous results obtained by analysing particular cases separately. We also show that these methods are robust with respect to different kinds of perturbations–which may come from computational errors, intentional deviations, as well as regularisation or approximation schemes–under surprisingly weak assumptions on the magnitude of the perturbations. Although we mostly focus on theoretical aspects, a numerical illustration based on the image inpainting problem reveals possible trends in the behaviour of these types of methods.

Introduction

Let $\mathcal{H}$ be a real Hilbert space. Krasnoselskii-Mann iterations approximate fixed points of a (quasi-)nonexpansive operator $T : \mathcal{H} \to \mathcal{H}$, by means of the update rule

$$x_{k+1} = (1-\lambda_k)x_k + \lambda_k T x_k,$$

where $\lambda_k \in (0,1)$ is a *relaxation* parameter.

The purpose of this work is threefold:

1. Provide a systematic and unified analysis of the convergence of the sequences produced by means of addition of Nesterov acceleration [2] and Polyak's heavy-ball momentum [3],
2. Establish the extent to which these iterations are stable with respect to perturbations,
3. Account for *diagonal* algorithms represented by a sequence of operators.

To this end, we study the behaviour of sequences generated iteratively by the set of rules

$$\begin{cases} y_k &= x_k + \alpha_k(x_k - x_{k-1}) + \varepsilon_k \\ z_k &= x_k + \beta_k(x_k - x_{k-1}) + \rho_k \\ x_{k+1} &= (1-\lambda_k)y_k + \lambda_k T_k z_k + \theta_k, \end{cases} \tag{1}$$

where $T_k : \mathcal{H} \to \mathcal{H}$, $\alpha_k, \beta_k, \lambda_k \in [0,1]$, and $\varepsilon_k, \rho_k, \theta_k \in \mathcal{H}$ for $k \geq 1$, and $x_0, x_1 \in \mathcal{H}$.

Convergence Analysis

Weak Convergence

We say a family (T_k) of operators is *asymptotically demiclosed* (at 0) if, for every sequence (u_k) in $\mathcal{H}$, such that $u_k \rightharpoonup u$ and $T_k u_k - u_k \to 0$, it follows that $u \in \bigcap_{k\geq 1} \mathrm{Fix}(T_k)$.

Theorem 1 *Let $T_k\colon \mathcal{H} \to \mathcal{H}$ be a family of quasi-nonexpansive operators such that*

$$F \coloneqq \bigcap_{k\geq 1} \mathit{Fix}(T_k) \neq \emptyset.$$

Suppose suitable conditions on the parameters hold, and assume the error sequences (ε_k), (ρ_k), and (θ_k) belong to $\ell^1(\mathcal{H})$. Assume, moreover, that $(I - T_k)$ is asymptotically demiclosed at 0. If (x_k, y_k, z_k) is generated by Algorithm (1)*, then (x_k, y_k, z_k) converges weakly to (p^*, p^*, p^*), with $p^* \in F$.*

Strong Convergence

Theorem 2 *Let $T_k\colon \mathcal{H} \to \mathcal{H}$ be a family of quasi-contractive operators such that $\mathit{Fix}(T_k) = \{p^*\}$. Suppose suitable conditions on the parameters hold, and assume the error sequences (ε_k), (ρ_k), and (θ_k) belong to $\ell^2(\mathcal{H})$. If (x_k, y_k, z_k) is generated by Algorithm* (1)*, then (x_k, y_k, z_k) converges strongly to (p^*, p^*, p^*). Moreover, $\sum_{k=1}^{\infty} \|x_k - p^*\|^2 < \infty$.*
If $\varepsilon_k \equiv \rho_k \equiv \theta_k \equiv 0$, then there exist $Q < 1$ and $C > 0$ such that

$$\|x_k - p^*\|^2 \leq C \cdot Q^k$$

for all $k \geq 1$.

Operators Not Sharing a Fixed Point

Our results can also be applied when $\bigcap_{k\geq 1} \mathrm{Fix}(T_k) = \emptyset$, if instead we are interested in the *Kuratowski lower limit* of the family $(\mathrm{Fix}(T_k))$. This set, which we denote by F_∞, consists of all $p_\infty \in \mathcal{H}$ for which there is a sequence (p_k), such that $p_k \in \mathrm{Fix}(T_k)$ for all $k \geq 1$, and $p_k \to p_\infty$. We say that (T_k) *nicely approximates* F_∞ if $u_k \rightharpoonup u$ and $T_k u_k - u_k \to 0$ together imply $u \in F_\infty$.

Based on Theorem 1, we obtain the following generalization:

Corollary 3 *Let $T_k\colon \mathcal{H} \to \mathcal{H}$ be a family of quasi-nonexpansive operators that nicely approximates $F_\infty \neq \emptyset$. Suppose suitable conditions on the parameters hold, and assume the error sequences (ε_k), (ρ_k), and (θ_k) belong to $\ell^1(\mathcal{H})$. Assume moreover that there exists a sequence (p_k) with $p_k \in \mathit{Fix}(T_k)$ and $(p_k - p_{k-1}) \in \ell^1(\mathcal{H})$. If (x_k, y_k, z_k) is generated by Algorithm* (1)*, then (x_k, y_k, z_k) converges weakly to some (p^*, p^*, p^*), with $p^* \in F_\infty$.*

Similar ideas allow us to formulate the following generalization of Theorem 2:

Corollary 4 *Let $T_k\colon \mathcal{H} \to \mathcal{H}$ be a family of quasi-contractive operators such that $\mathit{Fix}(T_k) = \{p_k\}$, with $p_k \to p^*$ and $(p_k - p_{k-1}) \in \ell^2(\mathcal{H})$. Suppose suitable conditions on the parameters hold, and assume the error sequences (ε_k), (ρ_k), and (θ_k) belong to $\ell^2(\mathcal{H})$. If (x_k, y_k, z_k) is generated by Algorithm* (1)*, then (x_k, y_k, z_k) converges strongly to (p^*, p^*, p^*). Moreover, $\sum_{k=1}^{\infty} \|x_k - p^*\|^2 < \infty$.*

Numerical Results

Consider the *image inpainting problem*: We represent an image X of M by N pixels by a tensor in $\mathcal{H} \coloneqq [0,1]^{M\times N\times 3}$, in which the three layers represent the red, green and blue colour channels. Denote by $\mathcal{A}$ the linear operator that maps an image to an image whose elements in the corrupted positions have been erased. The image inpainting problem is described by

$$\min_{X\in\mathcal{H}} \left\{ \frac{1}{2}\|X - X_{\text{corrupt}}\|^2 + \sigma\|X_{(1)}\|_* + \sigma\|X_{(2)}\|_* \right\},$$

where $\sigma > 0$ is a *regularisation parameter*, $X_{(1)} \coloneqq [X_{\cdot\cdot 1}\ X_{\cdot\cdot 2}\ X_{\cdot\cdot 3}]$, $X_{(2)} \coloneqq [X_{\cdot\cdot 1}^T\ X_{\cdot\cdot 2}^T\ X_{\cdot\cdot 3}^T]^T$ and $\|\cdot\|_*$ denotes the nuclear norm. This problem can be described by a fixed point problem of the family of operators (T_k) [1] given by

$$T_k \coloneqq I - J_{\rho_k B} + J_{\rho_k A} \circ (2J_{\rho_k B} - I - \rho_k C \circ J_{\rho_k B}),$$

where $A = \partial f$, $B = \partial g$, $C = \nabla(h \circ L)$, for $f(X) = \sigma\|X_{(1)}\|_*$, $g(X) = \sigma\|X_{(2)}\|_*$, $h(X) = \frac{1}{2}\|X - X_{\text{corrupt}}\|^2$ and $L = \mathcal{A}$, and can hence be solved by Algorithm (1).

The image to be inpainted has dimensions 512×512 pixels. We select a regularisation parameter of $\sigma = 0.5$ and corrupt the image randomly, with 50% of pixels erased. We always set $X_0 = X_1 = X_{\text{corrupt}}$.

We set $\rho_k \equiv \rho = 1.8$ and $\lambda_k \equiv \lambda = 0.8$, and add no perturbations (other than possible rounding errors by the machine). We run multiple versions of the algorithm, corresponding to different inertial schemes: no inertia, Nesterov, Heavy Ball, and reflected. In each case, we pick α and β such that the conditions in Theorem 1 are tight, and then select

$$\alpha_k = \left(1 - \frac{1}{k}\right)\alpha, \quad \beta_k = \left(1 - \frac{1}{k}\right)\beta.$$

The visual results are depicted in Figure 1. The convergence plots are presented in Figure 2

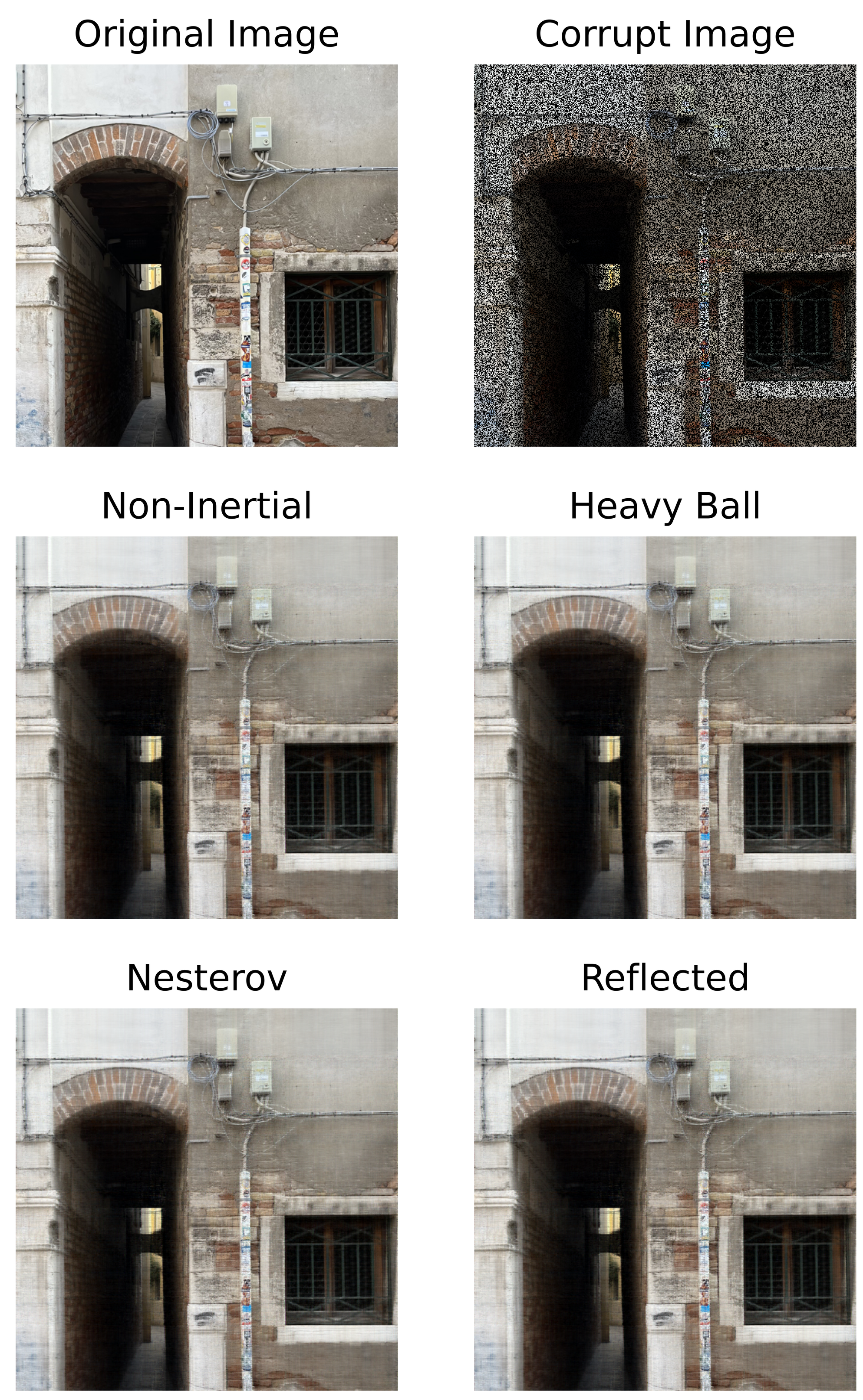

Figure 1: *Visual results for 50% corruption, with the parameters $\sigma = 0.5$, $\rho = 1.8$ and $\lambda = 0.8$.*

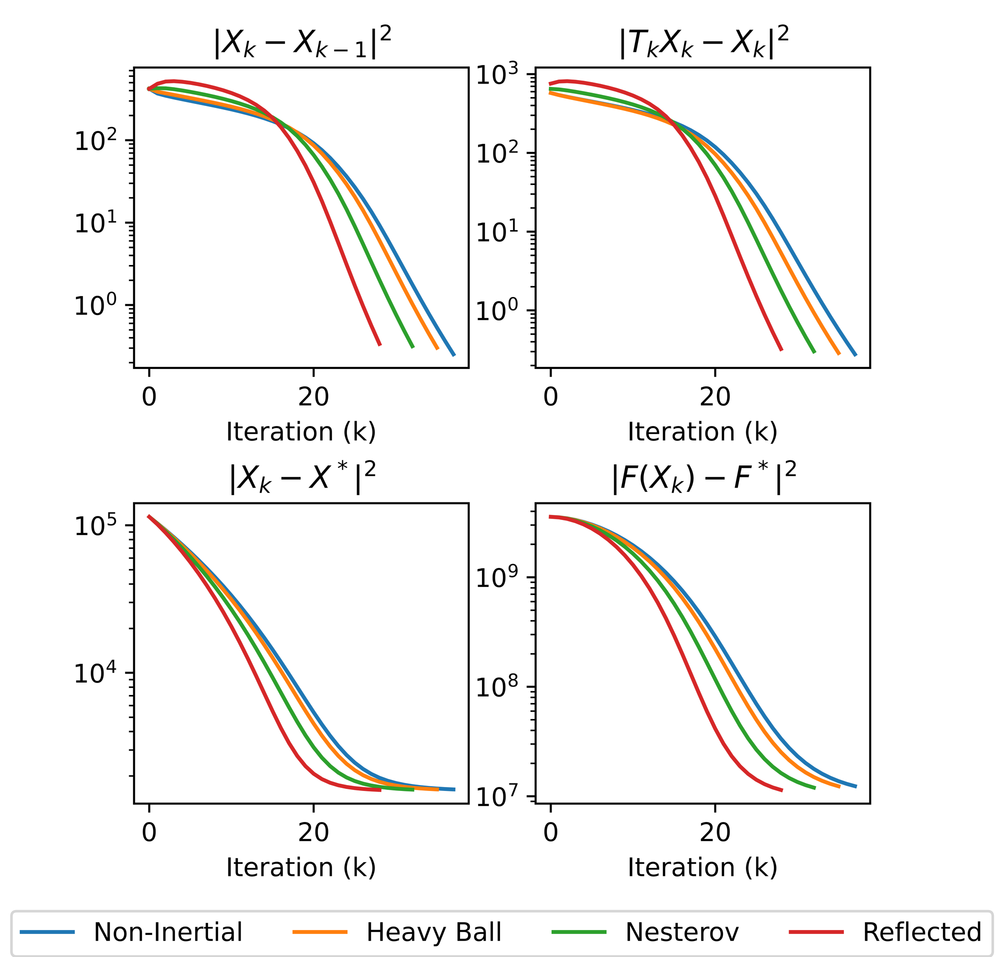

Figure 2: *Convergence plots for 50% corruption, with the parameters $\sigma = 0.5$, $\rho = 1.8$ and $\lambda = 0.8$.*

References

[1] D. Davis and W. Yin. A Three-Operator Splitting Scheme and its Optimization Applications. *Set-Valued and Variational Analysis*, 25:829–858, 2017.

[2] Y. Nesterov. A method for solving the convex programming problem with convergence rate $\mathcal{O}(1/k^2)$. *Proceedings of the USSR Academy of Sciences*, 269(3):543–547, 1983.

[3] B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964.

Download

Preprint

Poster