Renormalised steepest descent converges to a two-point attractor*

Luc Pronzato†, Henry P. Wynn‡ and Anatoly A. Zhigljavsky§

Abstract The behaviour of the standard steepest-descent algorithm for a quadratic function in $\mathbb{R}^d$ is investigated. We show that by rescaling the iterates to remain always on the unit sphere one can reveal special features of this behaviour. The renormalized algorithm is shown to converge to a two-point cycle on the unit circle. The cycle depends on the starting point in a complicated manner (the set of points converging to the same cycle is fractal), but all cycles belong to a particular plane, given by certain eigenvectors of the Hessian matrix of the objective function. The stability of the attractor is analysed. The rate of convergence of the algorithm is investigated. It is shown that the worst value of this rate is obtained only for some particular starting points. The introduction of a relaxation coefficient in the steepest-descent algorithm completely changes its behaviour, which may become chaotic. Different attractors are presented. We show that relaxation allows a significantly improved rate of convergence.

Key words Asymptotic behaviour; attractors; fractal; relaxation; renormalisation; steepest descent.

AMS math. subject class. 90C25, 68Q25, 58F12.

1 Introduction: an unsolved problem

For a general smooth function $f(\cdot)$ the steepest-descent algorithm is

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}), \qquad (1)$$

where

$$x^{(k)} = \left(x_1^{(k)}, \ldots, x_d^{(k)}\right)^T$$

*This work was supported by a UK Engineering and Physical Sciences Research Council for the second author and a grant of the French Ministère de l'Éducation Nationale (invitation triennale, procédure PAST) for the third author.

†Author for correspondence. Laboratoire I3S, CNRS–UNSA, bât. Euclide, " Les Algorithmes", 2000 route des Lucioles, Sophia Antipolis, 06410 Biot, France

‡Dept. of Statistics, University of Warwick, Coventry CV4 7AL, UK

§School of Mathematics, University of Wales, Senghennydd Road, Cardiff CF2 4YH, UK

is the k-th iterate of the algorithm,

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}\right)^T$$

is the gradient of $f(\cdot)$, and

$$\alpha_k = \arg\min_{\alpha} f(x^{(k)} - \alpha \nabla f(x^{(k)})) .$$

In [3] the authors studied the asymptotic behaviour of this algorithm for a quadratic function in $I\!R^d$:

$$f(x) = \frac{1}{2}(x - x^*)^T A (x - x^*) ,$$

where $x, x^* \in I\!R^d$ and A is a positive definite $d \times d$ matrix. This seemingly simple problem has some very complex aspects which are uncovered by the following renormalization idea. Define

$$v^{(k)} = \frac{x^{(k)} - x^*}{||x^{(k)} - x^*||} , \tag{2}$$

which has the effect of rescaling the iterates to remain always on the unit sphere: $||v|| = 1$.

The authors studied the behaviour of the process $\{v^{(k)}\}$ as $k \to \infty$. Except for pathological cases, in the limit the process is found to belong to the two dimensional space spanned by the eigenvectors (u_1, u_d) corresponding to the minimal and maximum eigenvalues λ_1 and λ_d of the matrix A under the assumption

$$0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_{d-1} < \lambda_d , \tag{3}$$

where λ_i is the i-th ordered eigenvalue of A. Ordering the $v_i^{(k)}$ compatibly with (3), the conjecture states that $v_i^{(k)} \to 0$ for $i = 2, \ldots, d-1$ as $k \to \infty$, and that the algorithm, in its renormalised form, converges to a two-point set on the circle $\{||v_1||^2 + ||v_d||^2 = 1\}$ in the plane spanned by u_1 and u_d. Numerical simulations show that the same convergence property holds for a convex differentiable function locally quadratic around its minimum x^*. It has been known for many years that the asymptotic convergence rate of the algorithm depends on the ratio $\rho = \lambda_d/\lambda_1$ (for instance, see [2], p. 152). Thus for any $x^{(k)} \in I\!R^d$

$$f(x^{(k+1)}) \leq \left(\frac{\rho - 1}{\rho + 1}\right)^2 f(x^{(k)}) \tag{4}$$

(see also Section 4). Despite an intensive search, no more precise properties concerning the rate of convergence were found in the literature. The authors doubts about the existence of results on the asymptotic behaviour of the steepest-descent algorithm were increased by the complexity of the behaviour of the renormalized algorithm en route to the limiting circle. This behaviour is fractal in nature, and it still remains open how the limiting asymptotic rate depends on the starting values. It is the worst-case rate, given by (4), not the actual rate, which simply depends on λ_1 and λ_d. However we are able to show that the algorithm attracts to two "conjugate" points on the circle. It is easy to check that *starting* on the circle the renormalized algorithm oscillates between the points, but without *attraction* to the circle this limiting behaviour is only a conjecture. We shall return to the discussion of these conjugate points and the corresponding asymptotic rate of convergence in Sections 3 and 4 respectively. The behaviour of the steepest-descent algorithm with relaxation is considered in Section 5.

2 Attraction theorem

Theorem 1 *Let the objective function f be*

$$f(x) = \frac{1}{2}(x - x^*)^T A(x - x^*)\,, \tag{5}$$

where A is a positive definite matrix with ordered eigenvalues $0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_{d-1} < \lambda_d$. Let $V = \mathrm{span}(u_1, u_d)$ be the two-dimensional plane generated by the (distinct orthogonal) eigenvectors u_1, u_d corresponding to λ_1 and λ_d, respectively. Then for any starting vector $x^{(1)}$ for which

$$u_1^T v^{(1)} \neq 0 \quad \text{and} \quad u_d^T v^{(1)} \neq 0 \tag{6}$$

the algorithm attracts to the plane V in the following sense:

$$w^T v^{(k)} \to 0 \quad k \to \infty$$

for any non-zero vector $w \in V^\perp$, where $v^{(k)}$ is defined by (2). Moreover, the sequence $\{v^{(k)}\}$ converges to a two-point cycle.

Proof.

For the quadratic function (5) the gradient of $f(\cdot)$ at $x^{(k)}$ equals

$$g^{(k)} = \nabla f(x^{(k)}) = A(x^{(k)} - x^*)\,,$$

and

$$\alpha_k = \frac{(g^{(k)})^T g^{(k)}}{(g^{(k)})^T A g^{(k)}}\,.$$

Therefore, the algorithm (1) can be rewritten as

$$x^{(k+1)} = x^{(k)} - \frac{(g^{(k)})^T g^{(k)}}{(g^{(k)})^T A g^{(k)}} g^{(k)}\,. \tag{7}$$

Without loss of generality we can assume that $x^* = 0$ and the matrix A is diagonal: $A = \mathrm{diag}(\lambda_1, \ldots \lambda_d)$ where $0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_{d-1} < \lambda_d$. With this assumption, the iteration of the algorithm (7) can be written as

$$x_i^{(k+1)} = x_i^{(k)} - \frac{\sum_{j=1}^d (g_j^{(k)})^2}{\sum_{j=1}^d \lambda_j (g_j^{(k)})^2} g_i^{(k)} \quad \text{for } \; i = 1, \ldots, d\,. \tag{8}$$

Multiplying both sides of i-th equation of (8) by λ_i we obtain

$$g_i^{(k+1)} = g_i^{(k)} - \frac{\sum_{j=1}^d (g_j^{(k)})^2}{\sum_{j=1}^d \lambda_j (g_j^{(k)})^2} \lambda_i g_i^{(k)} \quad \text{for } \; i = 1, \ldots, d\,. \tag{9}$$

Introduce now the variables

$$y_i^{(k)} = (g_i^{(k)})^2\,,$$

and their renormalised versions

$$z_i^{(k)} = \frac{y_i^{(k)}}{\sum_{j=1}^d y_j^{(k)}}\,. \tag{10}$$

Note that $z_i^{(k)} \geq 0$ for $i = 1, \ldots, d$ and $\sum_{j=1}^d z_j^{(k)} = 1$. The equations (9) then imply

$$y_i^{(k+1)} = \left(1 - \frac{\lambda_i \sum_{j=1}^d y_j^{(k)}}{\sum_{j=1}^d \lambda_j y_j^{(k)}} \right)^2 y_i^{(k)} \quad \text{for } \; i = 1, \ldots, d\,,$$

and

$$z_i^{(k+1)} = z_i^{(k)} \frac{\left(\sum_{j=1}^d \lambda_j z_j^{(k)} - \lambda_i \right)^2}{\sum_{l=1}^d \left(\sum_{j=1}^d \lambda_j z_j^{(k)} - \lambda_l \right)^2 z_l^{(k)}} \quad \text{for } \; i = 1, \ldots, d\,. \tag{11}$$

Note that the updating formula (11) is exactly the same for the weights $z_i^{(k)}$ and $z_j^{(k)}$ corresponding to equal $\lambda_i = \lambda_j$, and therefore these weights can be summed. It follows that a proof based on the assumption that all λ_i are different can easily be extended to the case of ties among $\lambda_2, \ldots, \lambda_{d-1}$. We thus assume that all eigenvalues of A are different: $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_d$. Note also that to prove the convergence to the two-dimensional plane V for the sequence $\{x^{(k)}/||x^{(k)}||\}$, it is enough to prove the same property for the sequence $\{z^{(k)}\}$, $k \to \infty$.

We divide the rest of the proof into four parts: we show **(i)** that the sequence (11) converges to a two-dimensional plane, **(ii)** that the directions are eventually fixed, and **(iii)** that they fix at u_1 and u_d. Finally, we show **(iv)** the convergence to a two-point cycle. It is convenient to consider the discrete measures

$$\pi_k = \left\{ \begin{array}{ccc} \lambda_1 & \ldots & \lambda_d \\ z_1^{(k)} & \ldots & z_d^{(k)} \end{array} \right\}$$

which place the weights $z_i^{(k)}$ at the points λ_i, respectively ($i = 1, \ldots, d$). These measures carry all information about the sequence (11). Let $\psi(\cdot)$ denote the operator corresponding to the application of (11) to a measure π, so that $\pi_{k+1} = \psi(\pi_k)$.

(i). Define the moments

$$\mu_m = \mu_m(\pi_k) = \int \lambda^m d\pi_k(\lambda) = \sum_{j=1}^d \lambda_j^m z_j^{(k)}\,, \; m = 0, \pm 1, \pm 2, \ldots \tag{12}$$

so that $\mu_0 = 1$, $\mu_1 = \sum_{j=1}^d \lambda_j z_j^{(k)}$, etc. Define also the moment matrices $M_n = M_n(\pi_k)$ by $\{M_n\}_{j,l} = \mu_{j+l-2}$ $(j, l = 1, \ldots, n)$ so that

$$M_2(\pi_k) = \begin{pmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{pmatrix} \quad \text{and} \quad M_3(\pi_k) = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 \\ \mu_1 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{pmatrix}.$$

Then the denominator in the right hand side of (11) is

$$D_k = \sum_{l=1}^d \left(\sum_{j=1}^d \lambda_j z_j^{(k)} - \lambda_l \right)^2 z_l^{(k)} = \mu_2 - \mu_1^2 = \det[M_2(\pi_k)]\,.$$

Using the updating formula (11) we have

$$\begin{aligned} D_{k+1} &= \sum_{l=1}^d \left(\sum_{j=1}^d \lambda_j z_j^{(k+1)} - \lambda_l \right)^2 z_l^{(k+1)} = \sum_{j=1}^d \lambda_j^2 z_j^{(k+1)} - \left(\sum_{j=1}^d \lambda_j z_j^{(k+1)} \right)^2 \\ &= \frac{1}{D_k}(\mu_1^2\mu_2 - 2\mu_1\mu_3 + \mu_4) - \frac{1}{D_k^2}(\mu_1^3 - 2\mu_1\mu_2 + \mu_3)^2\,. \end{aligned}$$

Therefore

$$D_{k+1} - D_k = \frac{2\mu_1\mu_2\mu_3 + \mu_2\mu_4 - \mu_1^2\mu_4 - \mu_3^2 - \mu_2^2}{(\mu_2 - \mu_1^2)^2} = \frac{\det[M_3(\pi_k)]}{\{\det[M_2(\pi_k)]\}^2}\,.$$

The conditions (6) transfer through the updating formula (11) to $z_1^{(k)} > 0$ and $z_d^{(k)} > 0$ for every $k = 1, 2, \ldots$ From simple moment theory we have therefore that the matrix $M_2(\pi_k)$ is positive definite for every $k = 1, 2, \ldots$ which implies $D_k = \det[M_2(\pi_k)] > 0$. Also, $M_3(\pi_k)$ is a non-negative definite moment matrix and $\det[M_3(\pi_k)] \geq 0$. Thus $D_{k+1} - D_k \geq 0$ which means that the sequence $\{D_k\}$ is monotonously non-decreasing. Since π_k is a probability measure with bounded support, $D_k = \det[M_2(\pi_k)]$ is bounded above by some constant D_* and therefore the sequence $\{D_k\}$ converges monotonically to a limit and $D_{k+1} - D_k \to 0$ when $k \to \infty$. Moreover,

$$\det[M_3(\pi_k)] = (D_{k+1} - D_k)D_k^2 \leq (D_{k+1} - D_k)D_*^2\,,$$

so that $\det[M_3(\pi_k)] \to 0$ when $k \to \infty$.

Note that using the Binet-Cauchy lemma

$$\begin{aligned} D_k &= \det[M_2(\pi_k)] = \sum_{i<j} z_i^{(k)} z_j^{(k)} (\lambda_j - \lambda_i)^2 \geq D_1 > 0\,, \\ \det[M_3(\pi_k)] &= \sum_{i<j<l} z_i^{(k)} z_j^{(k)} z_l^{(k)} (\lambda_j - \lambda_i)^2 (\lambda_l - \lambda_j)^2 (\lambda_i - \lambda_l)^2 \to 0, \quad k \to \infty\,. \end{aligned} \tag{13}$$

For a fixed iteration k define the pair (i_k, j_k) which achieves $\max_{i<j} z_i^{(k)} z_j^{(k)}$. (If there are several such pairs, take the smallest of them in, say, lexicographical order.) Then for every k we have

$$D_1 \leq D_k \leq z_{i_k}^{(k)} z_{j_k}^{(k)} \sum_{i<j} (\lambda_j - \lambda_i)^2\,.$$

Therefore

$$\delta \le z_{i_k}^{(k)} z_{j_k}^{(k)}\,, \quad \delta < z_{i_k}^{(k)} < 1-\delta\,, \quad \delta < z_{j_k}^{(k)} < 1-\delta\,, \tag{14}$$

where

$$\delta = \frac{D_1}{\sum_{i<j}(\lambda_j - \lambda_i)^2} > 0\,. \tag{15}$$

From (13) we have

$$\det[M_3(\pi_k)] \ge z_{i_k}^{(k)} z_{j_k}^{(k)} (\lambda_{j_k} - \lambda_{i_k})^2 \sum_{i \ne i_k, j_k} z_i^{(k)} (\lambda_i - \lambda_{i_k})^2 (\lambda_i - \lambda_{j_k})^2\,.$$

Since all λ_i are distinct, $\det[M_3(\pi_k)] \to 0$ and $z_{i_k}^{(k)} z_{j_k}^{(k)} \ge \delta > 0$, we have

$$\sum_{i \ne i_k, j_k} z_i^{(k)} \to 0, \quad k \to \infty\,. \tag{16}$$

This finishes part **(i)** of the proof. The interpretation is that although (i_k, j_k) depends on k the total weight associated with all other points tends to zero.

(ii). We will show in addition to (16), that there exists an iteration number k_* such that for all $k \ge k_*$ the pair (i_k, j_k) becomes fixed, that is, $(i_k, j_k) = (i_*, j_*)$ for some $1 \le i_* < j_* \le d$ and all $k \ge k_*$.

Let

$$\epsilon_* = \frac{\delta D_1}{(\lambda_d - \lambda_1)^2}\,,$$

where δ is defined in (15). According to (16), there exists $k_* = k_*(\epsilon_*) \ge 1$ such that the total weight associated with all other points different from (i_k, j_k) is smaller than ϵ_* for all $k \ge k_*$. Therefore, for $i \notin \{i_k, j_k\}$ and $k \ge k_*$ the updating formula (11) gives

$$z_i^{(k+1)} = z_i^{(k)} \frac{\left(\sum_{j=1}^d \lambda_j z_j^{(k)} - \lambda_i\right)^2}{D_k} < \frac{\epsilon_* (\lambda_d - \lambda_1)^2}{D_1} = \delta\,.$$

From (14), $z_{i_{k+1}}^{(k+1)} > \delta$ and $z_{j_{k+1}}^{(k+1)} > \delta$, and therefore $i \notin \{i_k, j_k\}$ implies $i \notin \{i_{k+1}, j_{k+1}\}$. This proves that $(i_k, j_k) = (i_*, j_*)$ for some $1 \le i_* < j_* \le d$ and all $k \ge k_*$.

(iii). Let us prove that $(i_*, j_*) = (1, d)$. Recall again that the assumption (6) and the updating formula (11) imply that $z_1^{(k)} > 0$ and $z_d^{(k)} > 0$ for all k.

Assume that $j_* < d$. Denote

$$\epsilon^* = \min\{\epsilon_*, \, \delta \min_{1 \le i < j < d} (\lambda_j - \lambda_i)/\lambda_d\}\,,$$

and assume that $k^* = k^*(\epsilon^*) \ge k_*$ is such that $(i_k, j_k) = (i_*, j_*)$ and the total weight associated with all other points different from (i_*, j_*) is smaller than ϵ^* for all $k \ge k^*$. The existence of such a k^* follows from (i) and (ii).

For convenience, rewrite the algorithm (11) in the form

$$z_i^{(k+1)} = z_i^{(k)} \frac{(\mu_1 - \lambda_i)^2}{D_k} \quad \text{for } \ i = 1, \dots, d\,, \tag{17}$$

and note that

$$\begin{aligned}
\mu_1 &= \sum_{j=1}^{d} \lambda_j z_j^{(k)} = \lambda_{i_*} z_{i_*}^{(k)} + \lambda_{j_*} z_{j_*}^{(k)} + \sum_{j \neq i_*, j_*} \lambda_j z_j^{(k)} \\
&\leq \lambda_{i_*} z_{i_*}^{(k)} + \lambda_{j_*} z_{j_*}^{(k)} + \lambda_d \sum_{j \neq i_*, j_*} z_j^{(k)} \leq \lambda_{i_*} z_{i_*}^{(k)} + \lambda_{j_*} z_{j_*}^{(k)} + \epsilon^* \lambda_d \\
&\leq \delta \lambda_{i_*} + (1-\delta) \lambda_{j_*} + \epsilon^* \lambda_d = \lambda_{j_*} + \delta(\lambda_{i_*} - \lambda_{j_*}) + \epsilon^* \lambda_d \\
&\leq \lambda_{i_*} - \delta \min_{1 \leq i < j < d} (\lambda_j - \lambda_i) + \epsilon^* \lambda_d \leq \lambda_{j_*} \,.
\end{aligned}$$

Therefore (17) implies that for all $k \geq k^*$

$$\frac{z_d^{(k+1)}}{z_d^{(k)}} = \frac{(\lambda_d - \mu_1)^2}{D_k} > \frac{(\lambda_{j_*} - \mu_1)^2}{D_k} = \frac{z_{j_*}^{(k+1)}}{z_{j_*}^{(k)}} \,.$$

We have arrived at a contradiction since the sequence $\{z_{j_*}^{(k)}\}$ is bounded from below by $\delta > 0$ while the sequence $\{z_d^{(k)}\}$ tends to zero. Thus, $j_* = d$ and analogously $i_* = 1$.

(iv). Finally, let D^* denote the limit $\lim_{k\to\infty} D_k$ discussed in **(i)**. There are only two discrete measures π^1, π^2 with nonzero weights on λ_1 and λ_d and such that

$$\det[M_2(\pi^1)] = \det[M_2(\pi^2)] = D^* \,,$$

namely

$$\pi^1 = \left\{ \begin{array}{cc} \lambda_1 & \lambda_d \\ p & 1-p \end{array} \right\} , \quad \pi^2 = \left\{ \begin{array}{cc} \lambda_1 & \lambda_d \\ 1-p & p \end{array} \right\} , \tag{18}$$

with

$$p = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{D^*}{(\lambda_d - \lambda_1)^2}} \,.$$

Note that $\psi(\pi^1) = \pi^2$ and $\psi(\pi^2) = \pi^1$. Therefore, convergence of D_k to D^* implies convergence of π_k to the limiting cycle $\pi^1 \to \pi^2 \to \pi^1 \to \cdots$

■

Remark 1 *Theorem 1 obviously generalizes to the case where (6) may not be satisfied. The algorithm then attracts to a two-dimensional plane V' and $\{v^{(k)}\}$ converges to a two-point cycle. V' is defined by the eigenvectors u_i, u_j associated with the smallest and largest eigenvalues such that $u_i^T v^{(1)} \neq 0$, $u_j^T v^{(1)} \neq 0$. For the sake of simplicity of notations, we shall assume in what follows that (6) is satisfied, and therefore $u_i = u_1$, $u_j = u_d$.*

Although the limiting behaviour of the algorithm is simple, its behaviour en route to the attractor is fairly complicated and presents a fractal structure. Figure 1 shows the projection onto the plane (v_1, v_3) of the region of attraction for $v^{(k)}$ to a small neighbourhood (radius < 0.02) of the point $v^* = (0.9606, 0, 0.2781)$, with $d = 3$, $\lambda_1 = 1$, $\lambda_2 = 2$,

$\lambda_3 = 4$. It illustrates the difficulty of predicting the limiting behaviour, and thus the aymptotic rate of convergence, see Section 4, as a function of the starting point.

Possible location of Figure 1

In Figure 2, the grey level of starting points on the unit sphere depends on the limiting value of p in (18). Again, this illustrates the complexity of the behaviour of the algorithm on the way to the attractor.

Possible location of Figure 2

3 Stability of attractors

From Theorem 1, the algorithm attracts to two conjugate points on the circle $\{||v_1||^2 + ||v_d||^2 = 1\}$, characterized by the discrete measure (18). However, some values of p^* correspond to unstable points. We shall use the following definition of stability, see [1] p. 444.

Definition 1 *A fixed point π for a mapping $T(\cdot)$ is called stable if $\forall \epsilon > 0$, $\exists \alpha > 0$ such that $\forall \pi^0$ for which $\|\pi^0 - \pi\| < \alpha$, $\|T^n(\pi^0) - \pi\| < \epsilon$ for all $n > 0$. A fixed point π is unstable if it is not stable.*

Consider a two-step iteration for $z_i^{(k)}$, $1 \le i \le d$:

$$\begin{aligned} z_i^{(k+2)} &= z_i^{(k+1)} \frac{(\sum_{j=1}^d \lambda_j z_j^{(k+1)} - \lambda_i)^2}{D_{k+1}} \\ &= z_i^{(k)} \frac{(\mu_1 - \lambda_i)^2}{D_k} \frac{1}{D_{k+1}} \left(\frac{\mu_1^3 - 2\mu_1\mu_2 + \mu_3}{D_k} - \lambda_i \right)^2 , \end{aligned} \tag{19}$$

and the corresponding transformation $\psi^2(\cdot)$ defined by $\pi_{k+2} = \psi^2(\pi_k)$. Since $\sum_{i=1}^d z_i^{(k)} = 1$ for all k, we substitute $1 - \sum_{i=1}^{d-1} z_i^{(k)}$ for $z_d^{(k)}$ in (19). This defines an operator $\phi(\cdot) : S_{d-1} \mapsto S_{d-1}$ on the $(d-1)$-dimensional canonical simplex

$$S_{d-1} = \{z = (z_1, \ldots, z_{d-1}) \mid z_i \ge 0, \sum_{i=1}^{d-1} z_i \le 1\} ,$$

which maps $(z_1^{(k)}, \ldots, z_{d-1}^{(k)})$ to $(z_1^{(k+2)}, \ldots, z_{d-1}^{(k+2)})$. Studying the properties of $\psi^2(\cdot)$ is equivalent to studying the properties of $\phi(\cdot)$.

Theorem 2
(i) Stability. *All points in the interval* $\mathcal{I}_S$ *defined by*

$$\mathcal{I}_S =]\frac{1}{2} - s(\lambda_{i^*}),\ \frac{1}{2} + s(\lambda_{i^*})[\,,$$

where

$$s(\lambda) = \frac{\sqrt{(\lambda_d - \lambda)^2 + (\lambda_1 - \lambda)^2}}{2(\lambda_d - \lambda_1)}\,,$$

and i^* *is such that* $|\lambda_{i^*} - \frac{\lambda_1 + \lambda_d}{2}|$ *is minimum over all* λ_i*'s,* $i = 2, \ldots, d-1$*, are stable.*
(ii) Instability. *All points in the set* $\mathcal{I}_U$ *defined by*

$$\mathcal{I}_U = [0,\ \frac{1}{2} - s(\lambda_{i^*})[\cup]\frac{1}{2} + s(\lambda_{i^*}),\ 1]$$

are unstable.

Proof.
(i) Stability. Take $p \in \mathcal{I}_S$ and consider the fixed point $z(p) = (p, 0, \ldots, 0)$ for $\phi(\cdot)$, $z(p) \in \mathcal{S}_{d-1}$. Assume that $z_i^{(k)} \leq \beta_1$, $i = 2, \ldots, d-1$ and $|z_1^{(k)} - p| \leq \beta_1$, that is $\|z^{(k)} - z(p)\|_\infty \leq \beta_1$. Then

$$z_i^{(k+2)} = z_i^{(k)} f(\lambda_i, p)[1 + O(\beta_1)]\,,\ \beta_1 \to 0\,, \tag{20}$$

where

$$f(\lambda, p) = \frac{(\mu_1^* - \lambda)^2[(\mu_1^*)^3 - 2\mu_1^*\mu_2^* + \mu_3^* - D(p)\lambda]^2}{[D(p)]^4}$$

with

$$\begin{aligned} \mu_m^* &= \lambda_1^m p + \lambda_d^m (1-p)\,,\ m = 1, 2, 3\,, \\ D(p) &= p(1-p)(\lambda_d - \lambda_1)^2\,. \end{aligned} \tag{21}$$

Then, $p \in \mathcal{I}_S$ implies $f(\lambda_i, p) < 1$, $i = 2, \ldots, d-1$. Define $f^*(p) = \max_{i \in \{2, \ldots, d-1\}} f(\lambda_i, p)$ and let K be any constant such that $f^*(p) < K < 1$. Then, from (20):

$$\exists \beta_0 \mid \forall \beta < \beta_0\,,\ z_i^{(k+2)} \leq K z_i^{(k)}\,,\ \forall i = 2, \ldots, d-1\,,$$

and thus

$$\forall \beta < \beta_0\,,\ z_i^{(k+2m)} \leq K^m z_i^{(k)}\,,\ \forall i = 2, \ldots, d-1\,,\ \forall m \geq 1\,. \tag{22}$$

Now, (13) implies $\det[M_3(\pi_k)] \leq C\beta_1$, with C some positive constant. Therefore,

$$|D_{k+1} - D_k| \leq \frac{C\beta_1}{D_k^2}\,. \tag{23}$$

Similarly, $\|z^{(k+1)} - z(1-p)\| \le \beta_2$ implies

$$|D_{k+2} - D_{k+1}| \le \frac{C\beta_2}{D_{k+1}^2} \le \frac{C\beta_2}{D_k^2}\,. \tag{24}$$

Since $\|z^{(k)} - z(p)\|_\infty \le \beta \Rightarrow \|z^{(k+1)} - z(1-p)\|_\infty \le H\beta$, with H some positive constant, (23,24) imply

$$|D_{k+1} - D_k| \le C'\beta\,,\ |D_{k+2} - D_{k+1}| \le C'\beta$$

when $\|z^{(k)} - z(p)\|_\infty \le \beta$, with $C' = (C/D_k^2)\max\{1, H\}$. This gives

$$|D_{k+2} - D_k| \le 2C'\beta\,.$$

Choosing $\beta < \beta_0$, one then obtain from (22)

$$|D_{k+2m+2} - D_{k+2m}| \le 2C'K^m\beta\,.$$

Therefore, $\forall l > 0$,

$$|D_{k+2l} - D_k| \le \frac{2C'\beta}{1-K} = O(\beta)\,,\ \beta \to 0\,.$$

Moreover, $\|z^{(k)} - z(p)\|_\infty \le \beta \Rightarrow |D_k - D(p)| \le A\beta$, with A a positive constant and $D(p)$ given by (21). Therefore, $\forall l > 0$,

$$|D_{k+2l} - D(p)| \le \left(\frac{2C'}{1-K} + A\right)\beta\,.$$

This, together with (22) implies

$$\forall l > 0\,,\ |z_1^{(k+2l)} - p| < L\beta\,,\ L > 1\,. \tag{25}$$

Finally, for any $\epsilon > 0$, define $\alpha = \min\{\epsilon/L, \beta_0/L\}$, (22) and (25) then imply the stability of the fixed point $z(p)$.

(ii) Instability. The Jacobian matrix J_ϕ of $\phi(\cdot)$ at points $z(p) = (p, 0, \ldots, 0)$ can be computed analytically, and is given by:

$$(J_\phi)_{ij} = \frac{\partial \phi_i(z)}{\partial z_j}_{|\mathrm{z=z(p)}} = \begin{cases} 1 & \text{if } i = j = 1 \\ \frac{(\lambda_j-\lambda_1)(\lambda_d-\lambda_j)^2[2\lambda_d(1-p)+2\lambda_1 p-\lambda_1-\lambda_j]}{p(1-p)^2(\lambda_d-\lambda_1)^4} & \text{if } j > 1 \text{ and } i = 1 \\ \frac{[\lambda_d(1-p)+\lambda_1 p-\lambda_j]^2[\lambda_d p+\lambda_1(1-p)-\lambda_j]^2}{p^2(1-p)^2(\lambda_d-\lambda_1)^4} & \text{if } i = j \\ 0 & \text{otherwise}\,. \end{cases}$$

One can easilly check that for $p \in \mathcal{I}_U$, $(J_\phi)_{ii} > 1$ for at least one i in $\{2, \ldots, d-1\}$. Since the $(J_\phi)_{ii}$'s are eigenvalues of J_ϕ, Theorem 15.16 in [1] indicates that $z(p)$ is unstable.

■

Note that $\forall \lambda_1, \ldots, \lambda_d$ the stability interval $\mathcal{I}_S$ contains the interval

$$]\frac{1}{2} - \frac{1}{2\sqrt{2}},\ \frac{1}{2} + \frac{1}{2\sqrt{2}}[\approx]0.14645,\ 0.85355[\,.$$

When $d = 3$, numerical simulations show that for any initial density of $x^{(1)}$ in $I\!R^d$ associated with a density of $z^{(1)}$ on S_{d-1} reasonably spread, the density of attractors $z(p)$ characterized by p can be approximated by the density

$$\varphi(p) = C \log \min\{1, (J_\phi)_{22}\} = \begin{cases} C \log(J_\phi)_{22} & \text{if } p \in \mathcal{I}_A \\ 0 & \text{otherwise}\,, \end{cases}$$

where C is a normalisation constant. Figure 3 shows the empirical density of attractors (full line) together with $\varphi(p)$ (dashed line) in the case $\lambda_1 = 1$, $\lambda_2 = 4$, $\lambda_3 = 10$. The support of this density coincides with the stability interval $\mathcal{I}_S$.

Possible location of Figure 3

When $d > 3$, the density of attractors depends on the initial density of $x^{(1)}$.

4 Rate of convergence

Define the convergence rate at iteration k by

$$r_k = \frac{f(x^{(k+1)})}{f(x^{(k)})}\,, \tag{26}$$

where $f(x)$ is given by (5). Without any loss of generality, one can take $f(x) = \sum_{i=1}^d \lambda_i x_i^2$. Rewriting r_k in terms of $z^{(k)}$ defined by (10), one gets

$$r_k = 1 - \frac{1}{\mu_1 \mu_{-1}}\,,$$

where μ_m is defined by (12). Define the asymptotic rate as

$$R = R(x^{(1)}, x^*) = \lim_{k \to \infty} \left(\prod_{j=1}^k r_j \right)^{1/k}\,. \tag{27}$$

Generally, R depends on the initial point $x^{(1)}$ and the optimal point x^*. Theorem 1 implies that for any fixed x^* and almost all $x^{(1)}$ the asymptotic rate depends only on the attractor (18) and is given by

$$R(p) = \left(\frac{f(x^{(k+2)})}{f(x^{(k)})} \right)^{1/2}$$

where $x^{(k)}$ corresponds to π^1 or π^2, see (18). This gives

$$R(p) = \frac{p(1-p)(\rho-1)^2}{[p+\rho(1-p)][(1-p)+\rho p]},$$

with $\rho = \lambda_d/\lambda_1$ the condition number of the matrix A. The function $R(p)$ is symmetric with respect to 1/2 and monotonously increasing from 0 to 1/2. The worst asymptotic rate is thus obtained at $p = 1/2$:

$$R_{\max} = \left(\frac{\rho-1}{\rho+1}\right)^2,$$

see (4). Note that from Kantorovich inequality, see [2], p. 151, $\mu_1\mu_{-1} \leq (1+\rho)^2/(4\rho)$, and therefore $\forall x^{(k)}$, $r_k \leq R_{\max}$. The worst rate is thus achieved only when $x_1^{(k)} = \pm\rho x_d^{(k)}$, $x_2^{(k)} = \cdots = x_{d-1}^{(k)} = 0$.

Consider now another convergence rate, defined by

$$R' = \lim_{k\to\infty} \left(\prod_{j=1}^{k} r'_j\right)^{1/k}.$$

where

$$r'_k = \frac{\|x^{(k+1)}\|^2}{\|x^{(k)}\|^2}.$$

Rewriting r'_k in terms of $z^{(k)}$, one gets

$$r'_k = 1 - \frac{2\mu_{-1}}{\mu_1\mu_{-2}} + \frac{1}{\mu_1^2\mu_{-2}}.$$

One can easilly check that for almost all $x^{(1)}$ the asymptotic rate R' is equal to $R(p)$, where p defines the attractor.

5 Steepest descent with relaxation

The introduction of a relaxation coefficient γ, with $0 < \gamma < 1$, in the steepest-descent algorithm totally changes its behaviour. The algorithm (7) then becomes

$$x^{(k+1)} = x^{(k)} - \gamma \frac{(g^{(k)})^T g^{(k)}}{(g^{(k)})^T A g^{(k)}} g^{(k)}.$$

For fixed A, depending on the value of γ, the renormalized process either attracts to periodic orbits (the same for almost all starting points) or exhibits a chaotic behaviour. Figures 4 (respectively 5) presents the classical period-doubling phenomenon in the case $d = 2$ when $\lambda_1 = 1$ and $\lambda_2 = 4$ (respectively $\lambda_2 = 10$). Figures 6 (respectively 7) give the asymptotic rate (27) as a function of γ in the same situation.

Possible location of Figure 4

Possible location of Figure 5

Possible location of Figure 6

Possible location of Figure 7

We get now instead of (26):

$$r_k(\gamma) = 1 - \frac{\gamma(2-\gamma)}{\mu_1\mu_{-1}}.$$

Note that from Kantorovich inequality the worst value of the rate is

$$1 - \gamma(2-\gamma)\frac{4\rho}{(1+\rho)^2} > R_{\max},$$

if $\gamma < 1$. However, numerical results show that for γ large enough the asymptotic rate is significantly better than $R_{\max}$. A detailed analysis of the 2-dimensional case gives the following results.

(i) If $0 < \gamma \leq \frac{2}{\rho+1}$, the process attracts to the fixed point $p = 1$ and $R = R(\gamma) = 1 - \frac{2\rho\gamma}{\rho+1}$.

(ii) If $\frac{2}{\rho+1} < \gamma \leq \frac{4\rho}{(\rho+1)^2}$, the process attracts to the fixed point $p = \frac{2\rho-\gamma(\rho+1)}{2(\rho-1)}$, and $R(\gamma) = R_{\max}$.

(iii) If $\leq \frac{4\rho}{(\rho+1)^2} < \gamma \leq \frac{2(\sqrt{2}+1)\rho}{(\rho+1)^2}$, the process attracts to the 2-point cycle (p_1, p_2), with

$$p_{1,2} = \frac{2\rho - \gamma(\rho+1) \pm \sqrt{\gamma(\gamma\rho^2 - 4\rho + 2\gamma\rho + \gamma)}}{2(\rho-1)},$$

and $R(\gamma) = 1 - \gamma$.

(iv) For larger values of γ one observes a classical period-doubling phenomenon, see Figures 4 and 5.

(v) If $\rho > 3 + 2\sqrt{2} \approx 5.828427$, the process attracts again to a 2-point cycle for values of γ larger than $\gamma_\rho = \frac{8\rho}{(\rho+1)^2}$, see Figure 5. For the limiting case $\gamma = \gamma_\rho$, the cycle is given by (p'_1, p'_2), with

$$p'_{1,2} = \frac{\rho\left(\rho^2 - 2\rho + 5 \pm 2\sqrt{(\rho^2 - 2\rho + 5)(5\rho^2 - 2\rho + 1)}\right)}{(\rho - 1)(\rho + 1)^3},$$

and the associated asymptotic rate is

$$R(\gamma_\rho) = \frac{\rho^2 - 6\rho + 1}{\rho^2 - 1}.$$

In higher dimensions, repeated numerical trials show that the process typically no longer attracts to the 2-dimensional plane spanned by (u_1, u_d). Figure 8 presents the projection of the attractor of $z^{(k)}$ on the plane (z_1, z_3) in the case $d = 3$, $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 4$ and $\gamma = 0.97$. Such a picture is typical for $d > 2$.

Possible location of Figure 8

References

[1] J. Hale and H. Koçak. *Dynamics and Bifurcations.* Springer-Verlag, Heidelberg, 1991.

[2] D.G. Luenberger. *Introduction to Linear and Nonlinear Programming.* Addison-Wesley, Reading, Massachusetts, 1973.

[3] L. Pronzato, H.P. Wynn, and A.A. Zhigljavsky. Dynamic systems in search and optimisation. *CWI Quarterly*, 8(3):201–236, 1995.

Figure captions

Figure 1. Projection of the region of attraction for $v^{(k)}$ to a small neighbourhood (radius < 0.02) of the point $v^* = (0.9606, 0, 0.2781)$ onto the plane (v_1, v_3) $(d = 3, \lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 4)$.

Figure 2. Starting points on the unit sphere colored a a function of the limiting value of p in (18) $(d = 3, \lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 4)$.

Figure 3. Empirical density of attractors (full line) and $\varphi(p)$ (dashed line) $(d = 3, \lambda_1 = 1, \lambda_2 = 4, \lambda_3 = 10)$.

Figure 4. Attractors for $z^{(1)}$ as a function of γ $(d = 2, \lambda_1 = 1, \lambda_2 = 4)$.

Figure 5. Attractors for $z^{(1)}$ as a function of γ $(d = 2, \lambda_1 = 1, \lambda_2 = 10)$.

Figure 6. Asymptotic rate (27) as a function of γ $(d = 2, \lambda_1 = 1, \lambda_2 = 4)$.

Figure 7. Asymptotic rate (27) as a function of γ $(d = 2, \lambda_1 = 1, \lambda_2 = 10)$.

Figure 8. Projection of the attractor of $z^{(k)}$ on the plane (z_1, z_3) $(d = 3, \lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 4, \gamma = 0.97)$.