

# Option Calcul Scientifique: Optimisation Numérique

Préparation Agrégation de Mathématiques  
Université de Rennes 1  
Isabelle Gruais

5 décembre 2023

## 1 Méthodes de gradient

Soit  $K \subset \mathbb{R}^N$  et soit  $J : K \rightarrow \mathbb{R}$ ,  $\alpha$ -convexe. Soit à résoudre : trouver  $x^* \in K$  solution de :

$$x^* \in K \quad \text{et} \quad J(x^*) = \min_{x \in K} J(x)$$

**Hypothèse 1.0.1.** *On suppose que  $J$  est  $\alpha$ -convexe, différentiable et que  $\nabla J$  est localement lipschitzienne.*

### 1.1 Gradient avec pas optimal

**Définition 1.1.1** (Algorithme du gradient avec pas optimal). On appelle suite engendrée par l'Algorithme du gradient avec pas optimal la suite  $(u_n)_{n \geq 0} \in (\mathbb{R}^N)^{\mathbb{N}}$  définie par la relation de récurrence :

$$u_0 \in \mathbb{R}^N \quad \text{arbitraire,}$$

$$u_{n+1} = u_n - \mu_n \nabla J(u_n)$$

avec

$$J(u_n - \mu_n \nabla J(u_n)) = \min_{\mu \in \mathbb{R}} J(u_n - \mu \nabla J(u_n)).$$

**Définition 1.1.2.** On appelle direction de descente à l'étape  $n$  le vecteur  $r_n := -\nabla J(u_n)$ .

**Proposition 1.1.1.** Dans l'algorithme 1.2.1, deux directions de descente consécutives sont orthogonales.

*Démonstration.* Soit  $n \geq 1$ . On pose :

$$f_n(\mu) = J(u_n + \mu r_n) \quad \text{où} \quad r_n := -\nabla J(u_n). \quad (1)$$

Par définition :

$$f'_n(\mu_n) = 0 = \underbrace{\langle \nabla J(u_n + \mu_n r_n), r_n \rangle}_{=-r_{n+1}} = -\langle r_{n+1}, r_n \rangle.$$

□

**Théorème 1.1.2.** Sous l'Hypothèse 1.0.1, la suite  $(u_n)_{n \geq 0}$  générée par l'algorithme du gradient avec pas optimal converge vers la solution  $u \in \mathbb{R}^N$  du problème de minimisation :

$$J(u) = \min_{v \in \mathbb{R}^N} J(v)$$

et on a l'estimation :

$$\|u_n - u\| \leq \frac{1}{\alpha} \|\nabla J(u_n)\|.$$

*Démonstration.* Par construction, la suite  $(J(u_n))_{n \geq 0}$  est décroissante, minorée par coercivité :

$$J(u_n) \geq \beta \|u_n\|^2 + b \geq b > -\infty$$

donc convergente. On en déduit qu'il existe  $C > 0$  t.q. :  $\forall n \geq 0$ ,

$$C \geq J(u_n) \geq \beta \|u_n\|^2 + b \Rightarrow \|u_n\|^2 \leq \frac{C}{\beta} =: M.$$

i.e. la suite  $(u_n)_{n \geq 0}$  est majorée.

Comme  $J$  est  $\alpha$ -convexe, on aussi :

$$\begin{aligned} J(u_n) &\geq J(u_{n+1}) + \langle \nabla J(u_{n+1}), u_n - u_{n+1} \rangle + \frac{\alpha}{2} \|u_n - u_{n+1}\|^2 = \\ &= J(u_{n+1}) + \underbrace{\mu_n \langle r_n, r_{n+1} \rangle}_{=0} + \frac{\alpha}{2} \|u_n - u_{n+1}\|^2 = J(u_{n+1}) + \frac{\alpha}{2} \|u_n - u_{n+1}\|^2 \end{aligned}$$

$$\Rightarrow \|u_n - u_{n+1}\|^2 \leq \frac{2}{\alpha}(J(u_n) - J(u_{n+1})) \xrightarrow{n \rightarrow +\infty} 0$$

Il en résulte :

$$\begin{aligned} \|\nabla J(u_n)\|^2 &= \langle \nabla J(u_n), \nabla J(u_n) \rangle = \langle \nabla J(u_n), \nabla J(u_n) - \nabla J(u_{n+1}) \rangle \\ &\leq C_M \|\nabla J(u_n)\| \|u_n - u_{n+1}\| \end{aligned}$$

d'où :

$$\|\nabla J(u_n)\| \leq C_M \|u_n - u_{n+1}\| \xrightarrow{n \rightarrow +\infty} 0.$$

Par  $\alpha$ -convexité de  $J$  :

$$\begin{aligned} \alpha \|u_n - u\|^2 &\leq \langle \nabla J(u_n) - \nabla J(u), u_n - u \rangle = \langle \nabla J(u_n), u_n - u \rangle \leq \|\nabla J(u_n)\| \|u_n - u\| \\ \Rightarrow \|u_n - u\| &\leq \frac{1}{\alpha} \|\nabla J(u_n)\| \xrightarrow{n \rightarrow +\infty} 0. \end{aligned}$$

□

**Proposition 1.1.3** (Cas particulier d'une fonctionnelle quadratique). *On suppose que  $J$  est définie par :*

$$J(x) = \frac{1}{2} \langle Ax, x \rangle - \langle b, x \rangle, \quad \forall x \in \mathbb{R}^N$$

où  $A \in \mathbb{R}^{N \times N}$  est symétrique, définie positive, et qu'il existe  $\alpha > 0$  t.q.

$$\langle Ax, x \rangle \geq \alpha \|x\|^2, \quad \forall x \in \mathbb{R}^N.$$

Alors :

$$\mu_n = \frac{\|r_n\|^2}{\langle Ar_n, r_n \rangle}, \quad \forall n \geq 0. \quad (2)$$

*Démonstration.* Avec la définition (1) :

$$\begin{aligned} f'_n(\mu_n) = 0 &= \langle A(u_n + \mu_n r_n) - b, r_n \rangle = \underbrace{\langle Au_n - b, r_n \rangle}_{=-r_n} + \mu_n \langle Ar_n, r_n \rangle = \\ &= -\|r_n\|^2 + \mu_n \langle Ar_n, r_n \rangle \end{aligned}$$

□

**Théorème 1.1.4.** *Sous les hypothèses de la Proposition 1.1.3, on note  $0 < \lambda_1 \leq \dots \leq \lambda_N$  la suite des valeurs propres de  $A$  comptées avec leur ordre de multiplicité. Alors, l'erreur à l'étape  $n$  définie par*

$$e_n := u_n - u$$

vérifie :

$$\frac{\langle Ae_n, e_n \rangle}{\langle Ae_0, e_0 \rangle} \leq \left( \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \right)^{2n}, \quad \forall n \geq 0$$

*Démonstration.* Le démonstration utilise le résultat suivant qui sera admis :

**Lemme 1.1.5** (Estimation de Kantorovitch).

$$\|x\|^4 \leq \langle Ax, x \rangle \langle A^{-1}x, x \rangle \leq \frac{1}{4} \left( \sqrt{c} + \frac{1}{\sqrt{c}} \right)^2 \|x\|^4 \quad (3)$$

où

$$c := \frac{\lambda_N}{\lambda_1}$$

Soit  $n \geq 0$ . On a :

$$e_{n+1} = u_{n+1} - u = \mu_n r_n + e_n$$

donc

$$\begin{aligned} \langle Ae_{n+1}, e_{n+1} \rangle &= \mu_n \underbrace{\langle Ae_{n+1}, r_n \rangle}_{=-\langle r_{n+1}, r_n \rangle} + \langle Ae_{n+1}, e_n \rangle = \langle Ae_{n+1}, e_n \rangle \\ &= \langle Ae_n, e_n \rangle + \mu_n \langle Ar_n, e_n \rangle = \langle Ae_n, e_n \rangle - \mu_n \|r_n\|^2 \\ &\stackrel{(2)}{=} \langle Ae_n, e_n \rangle - \frac{\|r_n\|^4}{\langle Ar_n, r_n \rangle} \end{aligned}$$

avec :

$$\langle Ae_n, e_n \rangle = \langle r_n, A^{-1}r_n \rangle.$$

On en déduit :

$$\begin{aligned} \frac{\langle Ae_{n+1}, e_{n+1} \rangle}{\langle Ae_n, e_n \rangle} &= 1 - \frac{\|r_n\|^4}{\langle r_n, A^{-1}r_n \rangle \langle A \langle Ar_n, r_n \rangle} \\ &\stackrel{(3)}{\leq} 1 - 4 \left( \sqrt{c} + \frac{1}{\sqrt{c}} \right)^{-2} = \left( \frac{c-1}{c+1} \right)^2 = \left( \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \right)^2. \end{aligned}$$

Il en résulte :

$$\frac{\langle Ae_{n+1}, e_{n+1} \rangle}{\langle Ae_0, e_0 \rangle} = \prod_{k=0}^n \frac{\langle Ae_{k+1}, e_{k+1} \rangle}{\langle Ae_k, e_k \rangle} \leq \left( \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \right)^{2(n+1)}.$$

□

## 1.2 Gradient avec pas fixe

**Définition 1.2.1** (Algorithme du gradient avec pas fixe). On appelle suite engendrée par l'Algorithme du gradient avec pas fixe  $\mu \in \mathbb{R}$  la suite  $(u_n)_{n \geq 0} \in (\mathbb{R}^N)^{\mathbb{N}}$  définie par la relation de récurrence :

$$u_0 \in \mathbb{R}^N \quad \text{arbitraire,}$$

$$u_{n+1} = u_n - \mu \nabla J(u_n)$$

**Théorème 1.2.1.** *Sous l'hypothèse 1.0.1, la suite  $(u_n)_{n \geq 0}$  définie par l'Algorithme 1.2.1 converge vers la solution du problème de minimisation :*

$$J(u) = \min_{v \in \mathbb{R}^N} J(v)$$

pour tout  $\mu > 0$  suffisamment petit. En particulier, elle converge pour tout  $\mu \in ]0, \frac{2\alpha}{C^2}[$  où  $C > 0$  est la constante de Lipschitz de  $\nabla J$  sur la boule fermée  $B := B(u, \|u - u_0\|)$ .

*Démonstration.* Soit  $\mathcal{P}(n)$  la propriété :  $u_n \in B$ . Alors  $\mathcal{P}(0)$  est vraie par définition de  $B$ . On suppose  $\mathcal{P}(n)$  vraie.

De la relation :

$$e_{n+1} = u_{n+1} - u_n + e_n = -\mu \nabla J(u_n) + e_n = -\mu(\nabla J(u_n) - \nabla J(u)) + e_n$$

on déduit que :

$$\begin{aligned} \|e_{n+1}\|^2 &= \|e_n\|^2 - 2\mu \langle \nabla J(u_n) - \nabla J(u), e_n \rangle + \mu^2 \|\nabla J(u_n) - \nabla J(u)\|^2 \\ &\leq \|e_n\|^2 \underbrace{(1 - 2\mu\alpha + C^2\mu^2)}_{=: \theta(\mu)} \end{aligned}$$

avec :  $\forall u, v \in \mathbb{R}^N$ ,

$$\alpha \|u - v\|^2 \leq \langle \nabla J(u) - \nabla J(v), u - v \rangle \leq C\alpha \|u - v\|^2$$

donc  $0 < \alpha < C$ . L'étude des variations de  $\theta : \mu \mapsto 1 - 2\mu\alpha + C^2\mu^2$  montre que

$$0 < \theta(\mu) < 1 \iff \mu \in ]0, \frac{2\alpha}{C^2}[.$$

On en déduit que si  $\mu \in ]0, \frac{2\alpha}{C^2}[$  alors

$$\|e_{n+1}\| < \|e_n\| \underset{\mathcal{P}(n)}{<} \|u - u_0\|$$

i.e.  $u_{n+1} \in B$ . Par récurrence sur  $n \geq 0$  :

$$\|e_n\| \leq \theta(\mu)^n \|e_0\| \xrightarrow{n \rightarrow +\infty} 0.$$

□

## Cas des fonctionnelles quadratiques

Soit  $J(x) = \frac{1}{2}\langle Ax, x \rangle$ ,  $\forall x \in \mathbb{R}^N$ . avec  $A$  symétrique définie positive de valeurs propres :

$$0 < \lambda_1 \leq \dots \leq \lambda_N.$$

Après diagonalisation dans une bon de vecteurs propres,  $J$  se réécrit sous la forme :

$$J(x) = \frac{1}{2} \sum_{i=1}^N \lambda_i x_i^2.$$

Soit  $\mu > 0$ . L'algorithme de gradient avec pas fixe  $\mu$  se réécrit :

$$u^0 \in \mathbb{R}^N, \quad u_i^{n+1} = (1 - \mu\lambda_i)u_i^n, \quad i = 1, \dots, N.$$

La suite  $(u^n)_{n \geq 0}$  converge ssi :  $\forall i \in [[1, N]]$ ,

$$|1 - \mu\lambda_i| < 1 \iff 0 < \mu < \frac{2}{\lambda_i}$$

i.e. ssi  $0 < \mu < \frac{2}{\lambda_N}$ . Le taux de convergence est optimal si  $\mu = \mu_{\text{opt}}$  est solution de

$$\max_{1 \leq i \leq N} |1 - \mu\lambda_i|$$

ce qui est réalisé pour

$$1 - \mu\lambda_1 = -1 + \mu\lambda_N \iff \mu = \frac{2}{\lambda_1 + \lambda_N} =: \mu_{\text{opt}}$$

## 1.3 Gradient conjugué pour une fonctionnelle quadratique

Soit  $J(x) = \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle$  où  $A$  est symétrique réelle, définie positive. On construit une suite  $(u_n)_{n \geq 0}$  t.q. :

$$u_{n+1} = u_n + \mu_n d_n, \quad n \geq 0$$

avec  $\mu_n \in \mathbb{R}$  solution du problème

$$J(u_{n+1}) = \min_{\mu \in \mathbb{R}} J(u_n + \mu d_n) \Rightarrow \langle \nabla J(u_n), d_n \rangle = 0$$

et où la suite de directions de descente  $(d_n)_{n \geq 0}$  est choisie t.q. :  $\langle Ad_n, d_{n-1} \rangle = 0$ ,  $n \geq 0$ .

**Définition 1.3.1** (Algorithme du gradient conjugué). On appelle suite engendrée par l'Algorithme du gradient avec pas fixe  $\mu \in \mathbb{R}$  la suite  $(u_n)_{n \geq 0} \in (\mathbb{R}^N)^{\mathbb{N}}$  définie par la relation de récurrence :

$$\begin{aligned} u_0 &\in \mathbb{R}^N \quad \text{arbitraire,} \\ d_0 &= -\nabla J(u_0) = b - Au_0 =: r_0 \\ r_n &= -\nabla J(u_n) = b - Au_n \\ d_n &= r_n + \beta_n d_{n-1}, \quad \text{où} \quad \beta_n = \frac{\langle r_n, Ad_{n-1} \rangle}{\langle d_{n-1}, Ad_{n-1} \rangle}, \\ u_{n+1} &= u_n + \alpha_n d_n, \quad \alpha_n = \frac{\langle r_n, d_n \rangle}{\langle d_n, Ad_n \rangle}, \end{aligned}$$

**Proposition 1.3.1.** *L'algorithme 1.3.1 vérifie :*

$$\begin{aligned} \langle r_{n+1}, d_n \rangle &= 0, \quad n \geq 0 \\ \langle Ad_n, d_{n-1} \rangle &= 0, \quad n \geq 1. \end{aligned}$$

*Démonstration.* Cela découle de la définition des coefficients  $\alpha_n, \beta_n, n \geq 0$ .  $\square$

**Proposition 1.3.2.** *L'algorithme 1.3.1 vérifie :*

$$\langle r_{n+1}, r_n \rangle = 0, \quad n \geq 0$$

*De plus, on a les expressions :*

$$\alpha_n = \frac{\|r_n\|^2}{\langle Ad_n, d_n \rangle}, \quad \beta_n = -\frac{\|r_n\|^2}{\|r_{n-1}\|^2}$$

*Démonstration.* Par construction :

$$\langle r_n, d_n \rangle = \|r_n\|^2 + \beta_n \underbrace{\langle r_n, d_{n-1} \rangle}_{=0} = \|r_n\|^2$$

et donc :

$$\alpha_n = \frac{\|r_n\|^2}{\langle Ad_n, d_n \rangle}.$$

Alors :

$$\begin{aligned} \langle r_{n+1}, r_n \rangle &= \|r_n\|^2 - \alpha_n \langle Ad_n, r_n \rangle = \\ &= \underbrace{\|r_n\|^2 - \alpha_n \langle Ad_n, d_n \rangle}_{=0} + \alpha_n \beta_n \underbrace{\langle Ad_n, d_{n-1} \rangle}_{=0} = 0 \end{aligned}$$

On a aussi :

$$\begin{aligned}
Ad_{n-1} &= \frac{1}{\alpha_{n-1}} A(u_n - u_{n-1}) = \frac{1}{\alpha_{n-1}} (r_{n-1} - r_n) \\
\Rightarrow \langle r_n, Ad_{n-1} \rangle &= \frac{1}{\alpha_{n-1}} \langle r_n, r_{n-1} - r_n \rangle = -\frac{\|r_n\|^2}{\alpha_{n-1}} \\
\langle d_{n-1}, Ad_{n-1} \rangle &= \frac{1}{\alpha_{n-1}} \langle d_{n-1}, r_{n-1} - r_n \rangle = \frac{1}{\alpha_{n-1}} \langle d_{n-1}, r_{n-1} \rangle \\
&= \frac{\|r_{n-1}\|^2}{\alpha_{n-1}} + \frac{\beta_{n-1}}{\alpha_{n-1}} \underbrace{\langle d_{n-2}, r_{n-1} \rangle}_{=0} = \frac{\|r_{n-1}\|^2}{\alpha_{n-1}}
\end{aligned}$$

d'où on déduit que

$$\beta_n = -\frac{\|r_n\|^2}{\|r_{n-1}\|^2}.$$

□

**Théorème 1.3.3.** *On pose :  $e_n = u_n - u$ ,  $\forall n \geq 0$ . Alors :*

$$\frac{\langle Ae_n, e_n \rangle}{\langle Ae_0, e_0 \rangle} \leq \left( \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \right)^{2n}, \quad \forall n \geq 0$$

*Démonstration.* Soit  $n \geq 0$ . De la relation  $e_{n+1} = \alpha_n d_n + e_n$ , on déduit :

$$\begin{aligned}
\langle Ae_{n+1}, e_{n+1} \rangle &= \langle Ae_n, e_n \rangle + 2\alpha_n \langle Ae_n, d_n \rangle + \alpha_n^2 \langle Ad_n, d_n \rangle = \\
&= \langle Ae_n, e_n \rangle - 2\alpha_n \langle r_n, d_n \rangle + \alpha_n^2 \langle Ad_n, d_n \rangle
\end{aligned}$$

avec

$$\alpha_n = \frac{\langle r_n, d_n \rangle}{\langle Ad_n, d_n \rangle}.$$

On en déduit :

$$\langle Ae_{n+1}, e_{n+1} \rangle = \langle Ae_n, e_n \rangle - \frac{\langle r_n, d_n \rangle^2}{\langle Ad_n, d_n \rangle} = \langle Ae_n, e_n \rangle \left( 1 - \frac{\langle r_n, d_n \rangle^2}{\langle Ad_n, d_n \rangle \langle Ae_n, e_n \rangle} \right)$$

avec  $\langle Ae_n, e_n \rangle = \langle r_n, A^{-1}r_n \rangle$ .

De plus, compte tenu de l'expression de  $\beta_n$  :

$$\begin{aligned}
\langle Ad_n, d_n \rangle &= \langle Ar_n, r_n \rangle + 2\beta_n \langle r_n, Ad_{n-1} \rangle + \beta_n^2 \langle d_{n-1}, Ad_{n-1} \rangle = \\
&= \langle Ar_n, r_n \rangle - \frac{\langle r_n, Ad_{n-1} \rangle^2}{\langle d_{n-1}, Ad_{n-1} \rangle} \in ]0, \langle Ar_n, r_n \rangle[.
\end{aligned}$$

Il en résulte, par croissance de  $x \mapsto -\frac{1}{x}$  sur  $]0, +\infty[$  :

$$\langle Ae_{n+1}, e_{n+1} \rangle \leq \langle Ae_n, e_n \rangle \left( 1 - \frac{\langle r_n, d_n \rangle^2}{\langle Ar_n, r_n \rangle \langle r_n, A^{-1}r_n \rangle} \right)$$

et on conclut comme pour le Théorème 1.1.4. □