219. EXTREMUMS : EXISTENCE, CARACTÉRISATION, RECHERCHE. EXEMPLES ET APPLICATIONS

I / Définitions et cadre de travail

__Cadre__ : Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n muni de la norme euclidienne $\|\ \|$. Soit $U \subset E$ non vide.

[R] p370 __déf 1.1__ Soit $a \in U$ et $f : U \longrightarrow \mathbb{R}$. On dit que f admet :

- un minimum global en a si $\forall x \in U,\ f(x) \geq f(a)$
- un minimum local en a s'il existe un voisinage V de a dans E tel que $\forall x \in V \cap U,\ f(x) \geq f(a)$
- un minimum strict global (resp local) si l'inégalité du 1er point (resp. 2ème) est stricte pour $x \neq a$.

__rq 1.2__ . La notion de maximum se définit en inversant le sens des inégalités
. Un extremum est un maximum ou un minimum

__ex 1.3__ $f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto \sin(x) \end{cases}$ admet un minimum global non strict (qui est aussi un minimum local strict) en $-\frac{\pi}{2}$

II / Compacité et existence d'extremums.

[T] p44 __Th 2.1__ Si $f : U \longrightarrow \mathbb{R}$ est continue et que U est un compact de E alors f est bornée sur U et atteint ses bornes

__ex 2.2__ Soit $N : (\mathbb{R}^n, \|\ \|_\infty) \longrightarrow \mathbb{R}$ une norme. N est continue sur le compact $S_\infty(0,1)$: elle y est donc bornée et elle y atteint son minimum

__app 2.3__ Équivalence des normes en dimension finie

[T] p46 __déf 2.4__ $f : U \longrightarrow \mathbb{R}$ est dite coercive si :

$$\forall A > 0\ \exists B > 0, \exists \alpha > 0, \forall x \in U, \left(\|x\| \geq B \text{ ou } d(x, U^c) \leq \alpha \right) \Longrightarrow f(x) \geq A$$

__rq 2.5__ Si $U = E$, f est coercive sur U ssi $\lim\limits_{\|x\| \to +\infty} f(x) = +\infty$

__Th 2.6__ Soit $f : U \longrightarrow \mathbb{R}$ continue et coercive. Alors f est minorée et atteint sa borne inférieure. [T] p46

__ex 2.7__ $f : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R} \\ (x,y) \longmapsto x^2 + y^2 \end{cases}$ est continue et coercive. Elle atteint son minimum en $(0,0)$

__app 2.8__ Théorème de d'Alembert : tout polynôme $P \in \mathbb{C}[X]$ non constant admet une racine complexe. [T] p48

III / Extremums et calcul différentiel

Dans cette partie, on suppose que U est __ouvert__

1) Ordre 1

__déf 3.1__ Soit $f : U \longrightarrow \mathbb{R}$ et $a \in U$ tel que $Df(a)$ existe. On dit que a est un point critique de f si $Df(a) = 0$

__Th 3.2__ Soit $f : U \longrightarrow \mathbb{R}$ et $a \in U$. Si a est un extremum local et si $Df(a)$ existe alors a est un point critique de f [R] p371

__rq - ex 3.3__ Cette condition est nécessaire pour avoir un extremum mais pas suffisante : $f : \begin{cases}]-1,1[\longrightarrow \mathbb{R} \\ x \longmapsto x^3 \end{cases}$ vérifie $Df(0) = 0$ mais n'a pas d'extremum en 0 [G1] p17

__app 3.4__ Théorème de Rolle : Soit $f : [a,b] \longrightarrow \mathbb{R}$, continue sur $[a,b]$, dérivable sur $]a,b[$ et telle que $f(a) = f(b)$ $(a < b)$. Alors $\exists c \in]a,b[,\ f'(c) = 0$ [G] p71

__app 3.5__ Théorème des accroissements finis : Soit $f : [a,b] \longrightarrow \mathbb{R}$ $(a < b)$ continue sur $[a,b]$ et dérivable sur $]a,b[$. Alors $\exists c \in]a,b[,\ f'(c) = \frac{f(b)-f(a)}{b-a}$ [G] p72

2) Ordre 2.

__Th 3.6__ Soit $f : U \longrightarrow \mathbb{R}$ et $a \in U$. Si a admet un minimum local en a et si $D^2 f(a)$ existe alors $Df(a) = 0$ et $D^2 f(a)$ est une forme quadratique positive. [R] p371

[DA] p18 ex 3.7 : $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto x^2+y^4$ admet un minimum en $(0,0)$ et on a bien $Df(0,0) = 0$ et

$$D^2f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \text{ est bien positive.}$$

[DA] p18 ct-ex 3.8 La condition 3.6 est nécessaire à l'existence d'un extremum mais pas suffisante : $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto x^2 - y^3$ admet $(0,0)$ comme point critique ; $D^2f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ est positive mais $(0,0)$ n'est pas un minimum local de f.

[R] p374 Th 3.9 Soit $f : U \longrightarrow \mathbb{R}$ et $a \in U$. Si $Df(a) = 0$ et $D^2f(a)$ est définie positive alors f admet un minimum local strict en a.

[R] p379 ex 3.10 $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto x^2 - y^2 + \frac{y^4}{4}$ admet un point critique en $(0,\sqrt{2})$ et $D^2f(0,\sqrt{2})$ est définie positive donc $(0,\sqrt{2})$ est un minimum local strict de f.

ct-ex 3.11 La condition 3.9 est suffisante à l'existence d'un extremum mais pas nécessaire : $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \longmapsto x^2+y^4$ admet un minimum local strict en $(0,0)$ mais $D^2f(0,0)$ n'est pas définie, comme vu en 3.7.

[G] p347 prop 3.12 Cas de la dimension 2. Soit $f : U \longrightarrow \mathbb{R}$ où $U \subset \mathbb{R}^2$ de classe C^2 sur U et $a \in U$ tel que $Df(a) = 0$.
- Si $\det(D^2f(a)) > 0$, a est un extremum local de f
- Si $\det(D^2f(a)) < 0$, f n'a pas d'extremum en a.
- Si $\det(D^2f(a)) = 0$, on ne peut rien dire sur a a priori

[G] p348 app 3.13 Principe du maximum harmonique
Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^2 telle que $\Delta f(x) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(x) = 0$ pour tout $x \in \mathring{B}(0,1)$. Alors $\forall x \in \mathring{B}(0,1)$,

$$\min_{\|y\|=1} f(y) \leq f(x) \leq \max_{\|y\|=1} f(y)$$

3) Problèmes sous contraintes

Th 3.14 Théorème des extrema liés [R] p373 et [G] p347
Soit $f, g_1, \dots, g_p : U \longrightarrow \mathbb{R}$ de classe C^1 et $\Gamma = \{x \in U,\ \forall 1 \leq i \leq p,\ g_i(x) = 0\}$.
Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$ et si les $Dg_i(a)$ sont linéairement indépendantes alors $\exists \lambda_1, \dots, \lambda_p \in \mathbb{R}$, nommés multiplicateurs de Lagrange, tels que $Df(a) = \sum_{i=1}^{p} \lambda_i Dg_i(a)$

Rq 3.15 Interprétation géométrique : $\nabla f(a)$ est orthogonal au plan tangent à Γ en a. [DA] p21

app 3.16 Inégalité arithmético-géométrique : [G] p349

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^+, \ \left(\prod_{i=1}^{n} x_i\right)^{1/n} \leq \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)$$

app 3.17 Diagonalisation des endomorphismes symétriques : [DA] p21
Soit $u \in \mathcal{L}(E)$ un endomorphisme symétrique.
Alors il existe une base orthonormée de E constituée de vecteurs propres de u.

IV / Extremums et convexité

1) Fonctions convexes

On suppose ici que U est convexe et non vide.

déf 4.1 Soit $f : U \longrightarrow \mathbb{R}$. On dit que f est convexe si [DA] p27
$\forall x,y \in U,\ \forall t \in [0,1],\ f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$
Si de plus l'inégalité est stricte pour $x \neq y$ et $t \in]0,1[$, on dit que f est strictement convexe.

Th 4.2 Si U est ouvert en plus d'être convexe et $f : U \longrightarrow \mathbb{R}$ est convexe on a : [T] p283
- si $a \in U$ est un maximum global alors f est constante
- si $a \in U$ est un minimum local de f alors ce minimum est global

[J] p253 Th 4.3 Si $f : U \to \mathbb{R}$ est convexe et différentiable sur l'ouvert convexe U alors f admet un extremum local en a si et seulement si $Df(a) = 0$

Ex 4.4 $f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto e^x \end{cases}$ est convexe et minorée mais comme $\forall x \in \mathbb{R}, Df(x) \neq 0$, f n'admet pas de minimum (ni local, ni global) sur $\mathbb{R}$

Th 4.5 Si $f : U \to \mathbb{R}$ est strictement convexe et admet un minimum alors il est unique

[X-ENS] p223 App 4.6 Ellipsoïde de John-Loewner (DEV 1)

Soit K un compact de $\mathbb{R}^n$ tel que $0 \in \mathring{K}$. Il existe un unique ellipsoïde centré en 0 et de volume minimal contenant K

2) Projection sur un convexe fermé

Th 4.7 Soit C un convexe fermé non vide de E et $x \in E$

Alors $\exists!\, P_C(x) \in C$ tel que $\|x - P_C(x)\| = d(x, C)$

et ce point est caractérisé par $\begin{cases} P_C(x) \in C \\ \forall y \in C, \langle x - P_C(x), y - P_C(x) \rangle \leq 0 \end{cases}$

[R] p884 App 4.8 Soit $m \in \mathbb{N}^*$ et $\forall i \in [\![1, m]\!]$, (x_i, y_i) un point de $\mathbb{R}^2$

On suppose que les x_i ne sont pas tous égaux. Alors $\exists!\, (\lambda, \mu) \in \mathbb{R}^2$ tq la somme $\sum_{i=1}^{m} (\lambda x_i + \mu - y_i)^2$ est minimale.

V Algorithmes de recherche et d'approximation

5.1 Descente de gradient à pas optimal. [HU] p53

Soit $f : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \frac{1}{2}\langle Ax | x \rangle + \langle b | x \rangle + c$

où $A \in S_n^{++}(\mathbb{R})$. Soit $x_0 \in \mathbb{R}^n$ et $\varepsilon > 0$

L'algorithme

tant que $\nabla f(x_k) \geq \varepsilon$

$$x_{k+1} = x_k - t_k \nabla f(x_k) \quad \text{où } t_k = \operatorname*{argmin}_{t \in \mathbb{R}} f(x_k - t\nabla f(x_k))$$

permet de construire la suite x_k qui tend vers $\bar{x} = \operatorname*{argmin}_{x \in \mathbb{R}^n} f(x)$ et telle que $\|x_k - \bar{x}\| \leq C \left(\frac{\lambda_1 - \lambda_n}{\lambda_n + \lambda_1} \right)^k$

où $C \in \mathbb{R}$ et $\lambda_n = \min Sp(A)$, $\lambda_1 = \max Sp(A)$

5.2 Méthode de Newton (DEV 2) [R] p152

Soit $f : [c, d] \to \mathbb{R}$ de classe C^2 telle que $f(c) < 0 < f(d)$ et $f'(x) > 0$ sur $[c, d]$.

La suite $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ converge vers a l'unique zéro de f sur $[c, d]$ si x_0 assez proche de a. De plus $|x_{n+1} - a| \leq C|x_n - a|^2$.

Si $f'' > 0$ alors $x_0 \in [a, d]$ suffit.

Corollaire 5.3 Si $g : [c, d] \to \mathbb{R}$ est de classe C^3 et strictement convexe, on peut approcher son minimum par la méthode de Newton appliquée à g'.

ANNEXE A

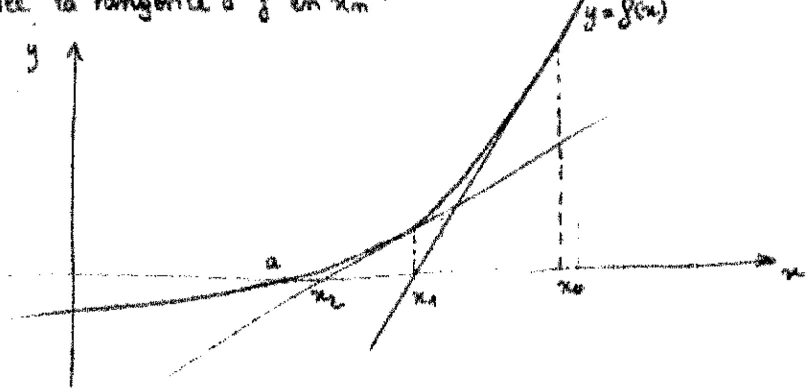
Méthode de descente dans le cas général :

$$x_{n+1} = x_n + \underbrace{t_n}_{\text{pas}} \underbrace{d_n}_{\text{direction de descente}}$$

Dans le cas d'une descente de gradient à pas optimal, $d_n = \nabla f(x_n)$ et le pas est optimal au sens où à chaque étape on minimise f sur la droite $x_n + \mathrm{Vect}(\nabla f(x_n))$.

ANNEXE B

Méthode de Newton : x_{n+1} est l'intersection de (Ox) avec la tangente à f en x_n.



Références

[R] = Rouvière, PGCD

[G] = Gourdon analyse

[OA] = Objectif agrégation

[X-ENS] = Oraux X-ENS algèbre 3

[HU] = Hiriart-Urruty, optimisation et analyse convexe

ELLIPSOÏDE DE JOHN-LOEWNER

<u>Th</u> Soit K un compact de $\mathbb{R}^n$ tel que $0 \in \mathring{K}$

Alors il existe un unique ellipsoïde contenant K, centré en 0 et de volume minimal

<u>démo</u> • Un ellipsoïde est un sous ensemble de $\mathbb{R}^n$ (que l'on munit ici de la norme euclidienne) de la forme

$$E_S = \{ X \in \mathbb{R}^n \mid {}^tXSX \leq 1 \} \text{ où } S \in S_n^{++}(\mathbb{R}) \}.$$

• Pour tout $S \in S_n^{++}(\mathbb{R})$, notons V_S le volume de E_S

Comme S est symétrique réelle, on peut la diagonaliser en base orthonormée donc supposer que $S = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$ où $a_i > 0$

On en déduit que
$$V_S = \int\!\!\cdots\!\!\int_{\sum_{i=1}^n a_i x_i^2 \leq 1} dx_1 \dots dx_n$$
$$= \int\!\!\cdots\!\!\int_{\sum_{i=1}^n t_i^2 \leq 1} \frac{dt_1 \dots dt_n}{\sqrt{a_1 \dots a_n}} \quad \text{via le changement de variable } x_i = \frac{t_i}{\sqrt{a_i}}$$
$$= \frac{1}{\sqrt{\det(S)}} V_{I_n}$$

• On peut donc reformuler le problème ainsi :

On cherche à minimiser $D \left\{ \begin{array}{l} S_n^{++}(\mathbb{R}) \longrightarrow \mathbb{R}^+ \\ S \longmapsto \dfrac{1}{\sqrt{\det(S)}} \end{array} \right.$

sur l'ensemble $A = \{ S \in S_n^{++}(\mathbb{R}) \,,\, K \subset E_S \}$

• Remarquons que $0 \in \mathring{K}$. Il existe donc $r > 0$ tel que $B(0,r) \subset K$. Comme $B(0,r) = E_{\frac{I_n}{r^2}}$, si $S \in A$ on a $E_{\frac{I_n}{r^2}} \subset K \subset E_S$ donc $V_{\frac{I_n}{r^2}} \leq V_S$ donc $D\left(\frac{I_n}{r^2}\right) \leq D(S)$

On en déduit que pour minimiser D sur A, il suffit de minimiser D sur $C = \{ S \in S_n^{++}(\mathbb{R}) \,,\, K \subset E_S \text{ et } D(S) \geq D\left(\frac{I_n}{r^2}\right) \}$

- Montrons que D est continue sur le compact non vide C.

  On en déduira que D atteint son minimum sur C d'où l'existence d'un ellipsoïde de volume minimal contenant K

- C est fermé car si $(S_n)_n \in C^{\mathbb{N}}$ converge vers S alors $S \in S_n^+(\mathbb{R})$ d'une part. D'autre part, en passant à la limite dans $\forall n \in \mathbb{N}$ $D(S_n) \geqslant D\left(\frac{I_n}{r^2}\right)$ on a $D(S) \geqslant D\left(\frac{I_n}{r^2}\right) > 0$ donc $S \in S_n^{++}(\mathbb{R})$

  Enfin, en passant à la limite dans $\forall X \in K$, $\forall n \in \mathbb{N}$, ${}^tXS_nX \leqslant 1$ on obtient que $\forall X \in K$, ${}^tXSX \leqslant 1$ donc $K \subset E_S$. Finalement, $K \subset E_S$

- C est borné. En effet, soit $S \in C$ et $X \in \mathbb{R}^n$ tel que $\|X\| \leqslant 1$

  Comme $rX \in B(0,r) \subset K \subset E_S$,

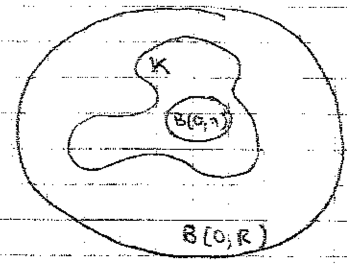
  $${}^t(rX)\,S(rX) = r^2\,{}^tXSX \leqslant 1 \text{ donc}$$

  $$|||S||| = \sup_{\|X\| \leqslant 1} {}^tXSX \leqslant \frac{1}{r^2}$$

- $C \neq \emptyset$ en effet, K est compact donc $\exists R > 0$, $K \subset B(0,R) = E_{\frac{I_n}{R^2}}$

  donc $\frac{I_n}{R^2} \in C$

- det et $\sqrt{\ }$ sont continues et $\sqrt{\det(S)}$ ne s'annule pas pour $S \in S_n^{++}(\mathbb{R})$ d'où la continuité de D sur C



- Montrons maintenant que D est strictement convexe sur le convexe C

  Cela assurera que le minimum trouvé au point précédent est unique.

- C est convexe : Soit $S, R \in C$ et $t \in [0,1]$

  $tS + (1-t)R \in S_n^{++}(\mathbb{R})$ par convexité de ce dernier

  De plus, $\forall X \in K$, ${}^tX(tS + (1-t)R)X = t\,\underbrace{{}^tXSX}_{\leqslant 1} + (1-t)\underbrace{{}^tXRX}_{\leqslant 1} \leqslant 1$

  Donc $K \subset E_{tS+(1-t)R}$

  Enfin, $B(0,r) \subset K \subset E_{tS+(1-t)R}$ donc $D(tS + (1-t)R) \geqslant D\left(\frac{I_n}{r^2}\right)$

- D est strictement convexe sur C (et sur $S_n^{++}(\mathbb{R})$ aussi d'ailleurs)

Soit $S \neq R \in C$ et $t \in]0,1[$

Par le théorème de pseudo réduction simultanée, il existe $P \in GL_n(\mathbb{R})$ tq

$$S = {}^tP \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix} P \quad \text{où } s_i > 0 \quad \text{et } R = {}^tPP$$

rq : Comme $S \neq R$, l'un des s_i est différent de 1. $(*)$

On a donc $D(tS + (1-t)R) = \det \left({}^tP \begin{pmatrix} ts_1 + 1-t & & 0 \\ & \ddots & \\ 0 & & ts_n + 1-t \end{pmatrix} P \right)^{-1/2}$

$$= \det(P)^{-1} \prod_{i=1}^{n} \underbrace{(ts_i + 1-t)^{-1/2}}_{= e^{-\frac{1}{2}\ln(ts_i + 1-t)}}$$

Par stricte concavité de $\ln$ sur $\mathbb{R}^{+*}$, $\forall i \in [\![1,n]\!]$,

$$\ln(ts_i + 1-t) \geq t\ln(s_i) + (1-t)\ln(1)$$

et l'une de ces inégalités est stricte par $(*)$

Donc $D(tS + (1-t)R) < \det(P)^{-1} \prod_{i=1}^{n} e^{-\frac{1}{2}(t\ln(s_i) + (1-t)\ln(1))}$

$$= \det(P)^{-1} e^{-\frac{1}{2}\left(t\ln\left(\prod_{i=1}^{n} s_i\right) + (1-t)\ln(1)\right)}$$

$$\leq \det(P)^{-1} \left(t e^{-\frac{1}{2}\ln(\prod s_i)} + (1-t) e^{-\frac{1}{2}\ln(1)} \right)$$

par convexité de $t \mapsto e^{-t/2}$ sur $\mathbb{R}$

$$= \det(P)^{-1} \left(t \left(\prod_{i=1}^{n} s_i\right)^{-1/2} + (1-t) \right)$$

$$= t\, D(S) + (1-t)\, D(R)$$

D'où le résultat annoncé

MÉTHODE DE NEWTON POUR L'APPROXIMATION DES ZÉROS

VLADISLAV TEMPEZ

Théorème: Méthode de Newton
Soit $f : [c,d] \to \mathbb{R}$ une fonction de classe C^2 telle que $f(c) < 0 < f(d)$ et $f'(x) > 0$.
f a alors un unique zéro $a \in [c,d]$ et la suite $x_{n+1} = x - \frac{f(x)}{f'(x)}$ converge vers a si x_0 est suffisamment proche de a. De plus il existe $C \in \mathbb{R}$ tel que

$$|x_{n+1} - a| \leq C|x_n - a|^2$$

Si $f'' > 0$ on a en plus que tout $x_0 \in [a,d]$ convient.

Démonstration :
f est continue sur $[c,d]$ et strictement croissante et $f(c) < 0 < f(d)$ donc f s'annule en un unique point $a \in]c,d[$.
Par ailleurs on pose $F(x) = x - \frac{f(x)}{f'(x)}$ et on a

$$F(a) = a - \frac{f(a)}{f'(a)}$$

et

$$F'(a) = 1 - \frac{f'(a)^2 - f(a)f''(a)}{f'(a)^2} = 1 - 1 = 0$$

De plus, $F(x) - a = x - a - \frac{f(x)-f(a)}{f'(x)}$ car $f(a) = 0$.
On obtient donc $F(x) - a = \frac{f(a)-f(x)-(a-x)'(x)}{f'(x)}$.
Or la formule de Taylor à l'ordre 2 appliquée en x donne

$$f(x) = f(a + x - a) = f(a) + (x-a)f'(x) + \frac{1}{2}(x-a)^2 f''(z)$$

avec $z \in]x,a[$.
On en tire donc

$$F(x) - a = \frac{f''(z)}{2f'(x)}(x-a)^2 \ (1)$$

En posant $C = \max_{x,z\in[c,d]} \frac{|f''(z)|}{2f'(x)}$ on a

$$|x_{n+1} - a| \leq C|x_n - a|^2$$

Pour b tel que $Cb < 1$ et $[a-b, a+b] \subset [c,d]$ on a $|x-a| < b \Rightarrow |F(x)-a| \leq Cb^2 < b$

donc $I =]a-b, a+b[$ est stable par F et si $x_0 \in I$ on a $C|x_{n+1}-a| \leq (C|x_n-a|)^2$ et donc par récurrence

$$|x_n - a| \leq \frac{(C|x_0-a|)^{2^n})}{C}$$

On a donc une convergence quadratique de x_n vers a car $C|x_0-a| \leq Cb < 1$.
Si de plus on a $f'' > 0$, on a par (1) que pour tout $x \in [c,d]$, $F(x) - a > 0$ et pour tout $x \in]a,d]$, $F(x) = x - \frac{f(x)}{f'(x)} < x$ car $f > 0$ au delà de a, donc $[a,d]$ est stable par F et si $x_0 \in I$, $(x_n)_n$ est décroissante et minorée donc converge vers un point fixe de F, c'est à dire a.
On tire aussi de (1) et du caractère C^2 de f que

$$\frac{x_{n+1}-a}{(x_n-a)^2} \sim_{n\to\infty} \frac{f''(a)}{2f'(a)}$$

(il suffit que $f''(a) > 0$ pour avoir l'équivalence).