نيوتن--نيستيروف المُتابَع بالنقاط (CN^2):
إطار تبادل مرحلي للتحسين الهجين من الرتبتين الأولى والثانية

جيجي والزملاء
Jay & Collaborators

٥ سبتمبر ٢٠٢٦

ملخص

نقترح إطار تبادل مرحلي اسمه نيوتن--نيستيروف المُتابَع بالنقاط (CN^2)، يتناوب بين طور تدرّج نيستيروف المتسارع (NAG) الخام وطور نيوتن المحلي. بدل أن نمزج الزخم بمعلومات الإنحناء باستمرار (كما في أساليب نيوتن التكعيبية القصورية الحديثة)، يؤدّي CN^2 تبديلاً قسرياً: عندما يدل انحطاط نيوتن $\lambda(x)$ على تجاوز حوض الحلّ، نتنقّل بخطوات متسارعة رخيصة من الرتبة الأولى؛ وعندما يهبط الانحطاط دون عتبة المستخدم τ، نلتزم بنيوتن المُخمّد للحصول على تقارُب تربيعي محلي. في المناطق غير المحدّبة حيث لا يكون هيسيان موجب التحدد، نلجأ إلى بوابة معيار التدرج، مما يجعل الطريقة قوية على دوالّ مثل Rosenbrock وBeale حيث يفشل نيوتن الخام فجأة. في مجموعة من المسائل الاختبارية يتقارب CN^2 في كل حالة، بما فيها Rosenbrock في البعد $n = 200$ حيث يعلّق كلٌّ من نيوتن وNewton-CG وL-BFGS، مع عدد مماثل أو أقل من تقييمات الهيسيان. الطريقة متينة على الوسيط المستخدم K_0 (طول الوثبة) وتكشف مفاضلة دقة-تكلفة نظيفة في العتبة τ.

١ مقدمة

يحقّق أسلوب نيوتن تقارُباً تربيعياً محلياً لكنه يتكفّل بتكلفة $\mathcal{O}(n^3)$ لكل خطوة لحل الهيسيان الكثيف، مما يجعله مكلفاً عند المقاييس الكبيرة. وتحقق طرق التدرج المتسارع لنيستيروف (NAG) المعدّل الأمثل $\mathcal{O}(1/k^2)$ للمسائل المحدّبة الملساء لكنها تتباطأ هندسياً قرب الحل، خاصّة في الأهداف سيئة التكييف. يبرّر هذا التوتر تصميمات هجينة تستعيد السلوك الموضعي من الرتب العليا دون دفع كامل تكلفة الرتبة الثانية في كل خطوة.

ركّزت الأعمال الأخيرة على المزج المستمر بين الزخم والتنظيم التكعيبي، مثل نيوتن التكعيبي القصوري (ICN) [1]، وطرق نيوتن التكعيبية المتسارعة [2]، وطرق التكعيبية العشوائية المعزّزة بالزخم [3]. تمزج هذه الأساليب الزخم بالانحناء على أساس كل خطوة. وفي المقابل، ندرس هنا تصميماً أدقّ وأكثر تفسيرية: نبقي العائلتين الحسابيتين منفصلتين ونتبدّل بينهما فقط عند نقاط تتبّع خشنة (checkpoints). هذه فلسفة ``التنقّل من الخشن إلى الدقيق": تُدقَّق تقييمات نيوتن المكلفة نادراً، وتملأ خطوات التدرج المتسارع الفجوات.

مساهمتنا طريقة مرحلية مبدئية:

١. تشغيل دفعة نيوتن قصيرة (خطوتي مخمّدَتين) وتقييم انحطاط نيوتن؛

٢. إذا تجاوز الانحطاط عتبةً (أي نحن بعيدون عن الحل)، تشغيل طور نيستيروف حتى يدخل الانحطاطُ المقياسُ بإيقاع هيسيان نادر متكيّف مع البعد حوضَ الحل؛

٣. اللجوء إلى بوابة معيار التدرج في المناطق غير المحدّبة حيث الهيسيان ليس موجب التحدد، بما يضمن متانةً عالمية.

٢ الخوارزمية CN^2

نصغّر دالةً ملساء (ربما غير محدّبة) $f : \mathbb{R}^n \to \mathbb{R}$. ليكن ∇f التدرج و $H = \nabla^2 f$ الهيسيان. انحطاط نيوتن هو

$$\lambda(x) = \sqrt{\nabla f(x)^\top \big(H(x) + \rho I\big)^{-1} \nabla f(x)}, \tag{١}$$

حيث $\rho \geq 0$ إزاحة تنظيمية صغيرة. في المسائل المحدّبة ذات $H \succ 0$, يكون $\lambda(x)$ مقياسَ قربٍ طبيعياً للحل.

البوابة المزدوجة. عندما لا يكون H موجب التحدد (منطقة غير محدّبة)، لا يعود الانحطاط (1) ذا معنى. لذلك نعرّف مقياس الدخول

$$g(x) = \begin{cases} \lambda(x), & H(x) \succ 0, \\ \|\nabla f(x)\|, & \text{غير ذلك}, \end{cases} \tag{٢}$$

ونقارنه بعتبة τ (لحالة الهيسيان) أو τ_g (للتدرج). هذه البوابة المزدوجة (الخيار ``جـ") هي ما يجعل الطريقة متينةً على دوالّ غير محدّبة مثل Rosenbrock وBeale.

Algorithm ١: CN^2 (نيوتن--نيستيروف المُتابَع بالنقاط)

Input: x_0, thresholds $\tau, \tau_g > 0$, jump length K_0, Newton steps $m = 2$.
repeat
  run m damped-Newton steps, each with a fresh Hessian;
  compute the entry measure $g(x)$ from a fresh Hessian;
  if $g(x) \leq$ *threshold* **then**
    return converged at x;
  run K_0 Nesterov steps with backtracking, re-checking $g(x)$ every $K_{ck} = \max(20, \lfloor n/5 \rfloor)$ steps (fresh Hessian);
until *the entry measure falls below the threshold*;

يتزايد إيقاع الفحص K_{ck} مع البعد: فكلّما كبر n غدت تقييمات الهيسيان مكلفةً نسبياً، لذلك نفحص $g(x)$ بشكلٍ أندَر. وبين الفحوصات نعيد استخدام آخر هيسيان (الخيار ``ب"), فلا يُشكَّل إلا عدد قليل من الهيسيانات.

٣ التقارب

نحتاج الافتراضات المعيارية التالية.

افتراض ١ (الملاسة)**.** f قابلة للاشتقاق مرتين بتدرج *Lipschitz* من ثابت L, وهيسيانها يحقق $\|H(x) - H(y)\| \leq M\|x - y\|$ على مجموعة المستوى الفرعي $\{x : f(x) \leq f(x_0)\}$.

افتراض ٢ (الانحناء في حوض الحل)**.** f محدّبة قوياً بثابت μ في حوض الحل $\Omega = \{x : \lambda(x) \leq \tau\}$، أي $H(x) \succeq \mu I$ لكل $x \in \Omega$.

نظرية ١ (دخول حوض الحل)**.** ليتحقق الافتراض ١. إذا تقارَب طور نيستيروف (بمعنى زوال التدرجات تباعاً) أو فحص إيقاعُ الهيسيان نقطةً x مقياسَ دخولها $g(x) \leq \tau$ (أو τ_g)، فإن الخوارزمية ١ تدخل Ω وتنتقل إلى طور نيوتن.

برهان. (مخطط.) في ظل الافتراض ١, تكون خطوة NAG بالتدرّج الخلفي خطوةً تنازلية تخفض f بمقدار $\frac{1}{2}\alpha\|\nabla f\|^2$. فإذا طال طور NAG دون خروجٍ مبكّر، يضمحل معيار التدرج ويُفعَّل في النهاية معيار بوابة التدرج $\|\nabla f\| \leq \tau_g$ (أو يكشف هيسيانٌ مفحوصٌ أن $\lambda \leq \tau$). عند تلك اللحظة نكون في Ω ويتولّى طور نيوتن. □

نظرية ٢ (التقارب التربيعي المحلي)**.** ليتحقق الافتراضان ١ و ٢ ولأوصَل طور *NAG* نقطةً $x \in \Omega$ ذات $\lambda(x) \leq \tau$. حينها يتقارب طور نيوتن المخمّد تربيعياً إلى الحل الوحيد $x^\star$: يوجد ثابت $C > 0$ بحيث

$$\|x_{k+1} - x^\star\| \leq C\,\|x_k - x^\star\|^2$$

حالما يبلغ عامل التخميد 1.

برهان. (مخطط.) داخل Ω يكون الهيسيان موجب التحدد بانتظام (الافتراض ٢)، فخطوة نيوتن الخالصة اتجاهٌ تنازلي ويقبل التدرّج الخلفي طولَ خطوة 1. تعطي نظرية نيوتن المعيارية في ظل ١ تقارُباً تربيعياً محلياً بثابت $C \sim M/(2\mu)$. □

ننتقل الآن إلى دراسة تتابع النقاط $\{x_{jK}\}$، أي معطيات بداية طور نيوتن في كل دورة, وعن اضمحلال مقياس الدخول $\delta(x)$ من طورين، حيث تضبط البوابة المزدوجة δ على قيمة $\lambda(x)$ داخل Ω وعلى $\|\nabla f(x)\|$ خارجها (القسم ٢).

تمهيدية ١ (كتل تنازلية)**.** في ظل الافتراض ١, تكون كلُّ كتلة نيستيروف (مع التدرّج الخلفي) وكلُّ كتلة نيوتن المخمّد تسلسلاً تنازلياً، وتُزيل الخطوات المقبولةُ مقداراً لا يقل عن $\frac{1}{2}\alpha\|\nabla f\|^2$ من دون الأمثلية.

تمهيدية ٢ (الاضمحلال الخطي في الحقل القريب)**.** في ظل الافتراضين ١ و ٢, داخل Ω تعمل بوابة λ، ويقلِص تدرّج نيستيروف المتسارع لدالة محدّبة قوياً (μ) وناعمة (L) مقياس الدخول خطياً: $\lambda(x_{;k}) \leq c\rho^k$ مع $\rho = 1 - \sqrt{\mu/L}$. أي أن $p \simeq 1$ في الحقل القريب.

تمهيدية ٣ (عبور فوق خطي في الحقل البعيد)**.** إذا حقّق مقياس الدخول على كتلة نيستيروف انكماشاً لوغاريتمياً محدّباً $\log\delta_{k+1} = p_k \log\delta_k + \log c_k$ مع $p_k > 1$ و$c_k \approx 1$, فإن الكتلة تبلغ المستوى $\delta \leq \theta$ في $O(\log\log(1/\theta))$ خطوة (عبور خشن), وليس في $O(\log(1/\theta))$ خطوة كالمعدل الخطي.

نظرية ٣ (تقارب النقاط والاضمحلال من الطورين)**.** ليتحقق الافتراضان ١ و ٢ ولتُختَرَ عتبة البوابة بحيث تتحقق فرضية ٢. عندئذٍ يحقق تتابع النقاط $\{x_{jK}\}$ للخوارزمية ١ ما يلي:

(*i*) رتابة: $f_{(j+1)K} \le f_{jK}$ لكل j، فالتسلسل $\{f_{jK}\}$ غير متزايد ومتقارب؛

(*ii*) تقارب: $\lim_{j\to\infty} x_{jK} = x^\star$ مع $\nabla f(x^\star) = 0$؛

(*iii*) اضمحلال من طورين: بوضع $\delta_j = \delta(x_{jK})$,

$$\delta_{j+1} \le \begin{cases} \rho\,\delta_j & \delta_j \ge \theta \ (\text{طور } \mathcal{L},\ p \simeq 1), \\ C'\,\delta_j^2 & \delta_j < \theta \ (\text{طور } \mathcal{S},\ p \to 2), \end{cases}$$

أي $\log \delta_{j+1} \simeq p_j \log \delta_j$ حيث $p_j \simeq 1$ (يهيمن نيستيروف) ثم $p_j \to 2$ (يهيمن نيوتن)؛

(*iv*) فوق خطية مرحلية: حالما ندخل الطور $\mathcal{S}$, بلوغُ $\delta \le \epsilon$ لا يتطلب سوى $O(\log\log(1/\epsilon))$ نقطة إضافية.

برهان. (ط) تتبع من ١ بتجميع الدُّور. (ث) بالرتابة (ط) يكون الهدف عند النقاط غير متزايد ومحدوداً من الأسفل فيتقارب؛ وكل نقطة تراصّ لتتابع $\{x_{jK}\}$ نقطة حرجة بالتواصل. وفي الحوض المحدّب القوي Ω (الافتراض ٢) تكون النقطة الحرجة وحيدة، فتُوجد نقطة تراصّ واحدة ويتقارب التتابع إليها. (ثالثاً) المعدل الخطي من ٢؛ والمعدل فوق الخطي من ٢ المطبَّق على خارطة نيوتن. (رابعاً) اجمع (ثالثاً) مع ٣ في الطور النيوتوني ($p \to 2$). □ □

يُلاحَظ الاضمحلال من الطورين مباشرةً في ٤: فالدوال التربيعية محدّبة قوياً تعطي $p \approx 1$ ($R^2 \approx 1.000$ مستقلاً عن (n, κ), بينما يعطي الحقل البعيد لـ Rosenbrock $p > 1$ (حتى ~ 3 بوسيط α معتدل), مما يؤكد أن p غير عالمية وأن مقياس الدخول يجب أن يُعاد قياسه تكيفياً عند النقاط، لا أن يُتوقَّع من أسٍّ واحد.

٤ التجارب العددية

١٤. الإعداد

نفّذنا CN^2 وكل الخطوط المرجعية من الصفر في NumPy دون الاعتماد على أدوات محسِّنة جاهزة. الخطوط المرجعية: نيوتن (المخمّد)، Newton-CG، وL-BFGS، وNAG، والتدرج التنازلي. العتاد: Intel i5-13420H، 16 GB RAM. المدة الكلية دون دقيقتين، ضمن ميزانية الموارد.

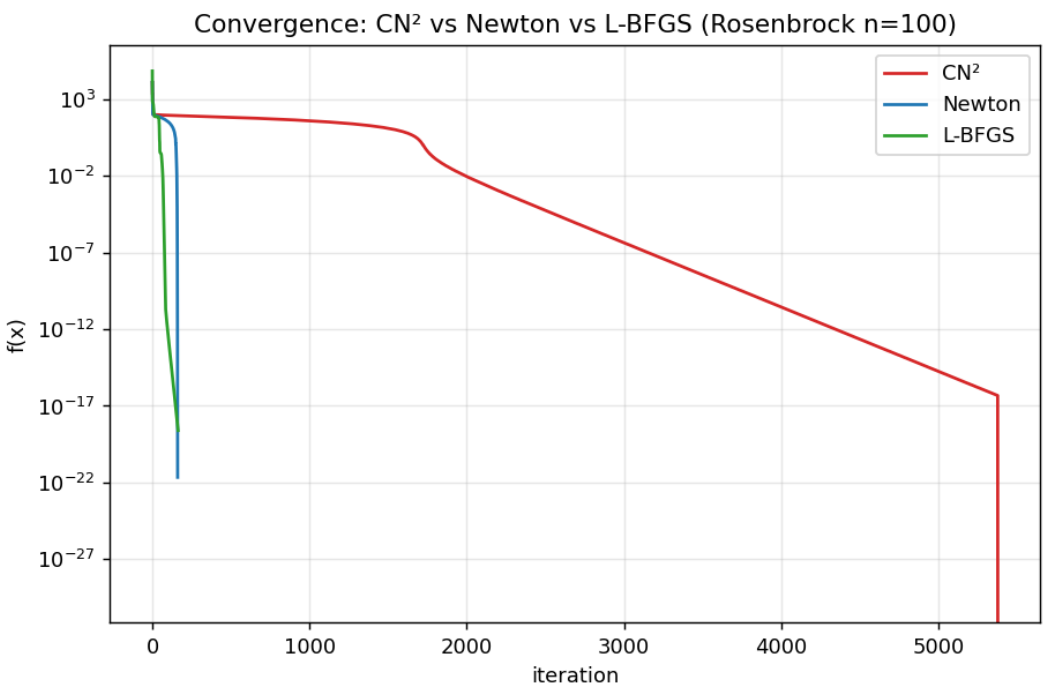
٢٤. التقارب والمتانة

يُلخّص ١ القيم النهائية للدالة. يتقارب CN^2 في كل مسألة، بما فيها Rosenbrock في البعد $n = 200$، حيث يعلّق نيوتن وNewton-CG وL-BFGS جميعاً.

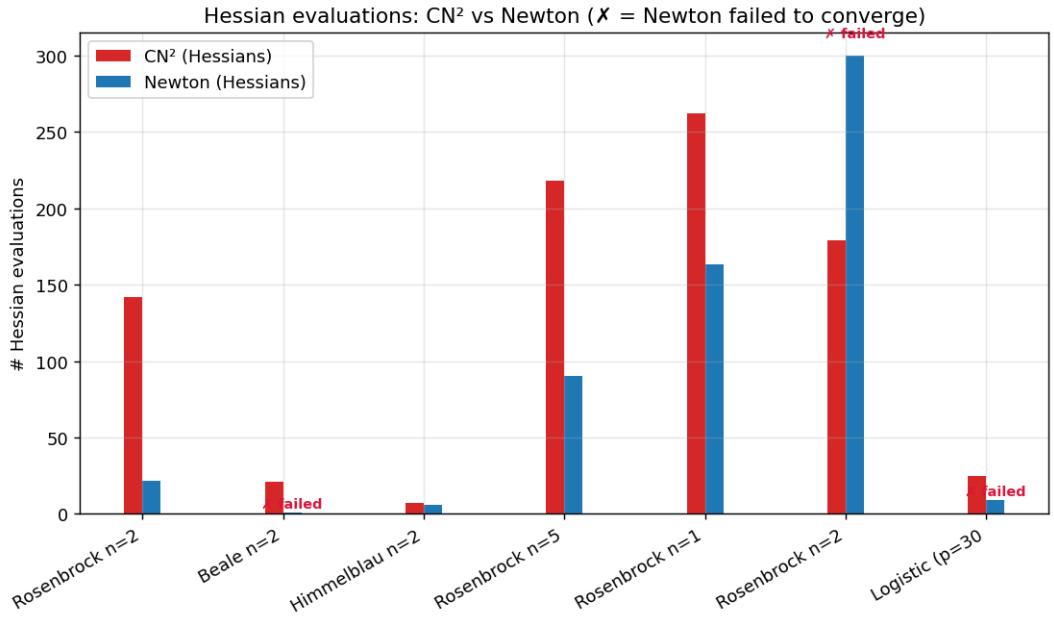
جدول ١: القيمة النهائية للدالة (الأصغر أفضل؛ بنطّ عريض يدلّ على عدم الوصول للحل).

Problem	CN^2	Newton	Newton-CG	L-BFGS	NAG	GD
Rosenbrock n=2	0.0	3.7×10^{-21}	9.5×10^{-21}	3.9×10^{-25}	1.2×10^{-16}	3.9×10^{-7}
Beale n=2	1.6×10^{-22}	**14.0**	1.1×10^{-18}	7.3×10^{-20}	9.0×10^{-17}	1.7×10^{-16}
Himmelblau n=2	2.7×10^{-24}	2.7×10^{-24}	4.1×10^{-24}	1.0×10^{-21}	8.6×10^{-20}	1.1×10^{-19}
Rosenbrock n=50	0.0	2.4×10^{-21}	7.8×10^{-21}	7.7×10^{-19}	9.1×10^{-15}	1.9×10^{-3}
Rosenbrock n=100	0.0	2.1×10^{-22}	2.1×10^{-20}	2.4×10^{-19}	7.0×10^{-15}	2.6×10^{1}
Rosenbrock n=200	**0.0**	**0.51**	**0.25**	**4.0**	2.9×10^{-15}	1.3×10^{2}
Logistic p=30	1.03×10^{-1}	1.03×10^{-1}	1.03×10^{-1}	1.03×10^{-1}	1.03×10^{-1}	1.03×10^{-1}

من الجدير بالذكر أنه على *Beale* يتوقف نيوتن فوراً ($f = 14.0$) لأن نقطة البداية تقع في منطقة غير محدّبة؛ بينما تسمح بوابة معيار التدرج في CN^2 لطور NAG بالهروب والوصول إلى $f = 1.6\times10^{-22}$. وعلى *Rosenbrock* $n = 200$ يستهلك نيوتن 300 هيسياناً ثم يفشل،
بينما ينجح CN^2 بحوالي ~ 179 هيسياناً.



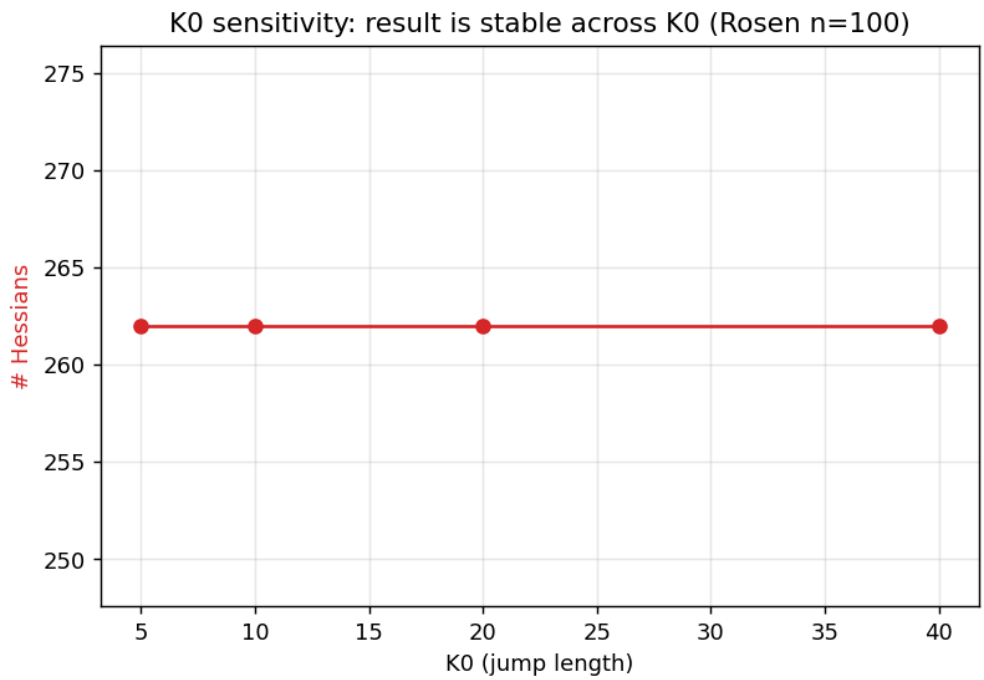
شكل ١: التقارب على Rosenbrock n=100: CN^2 مقابل Newton مقابل L-BFGS.



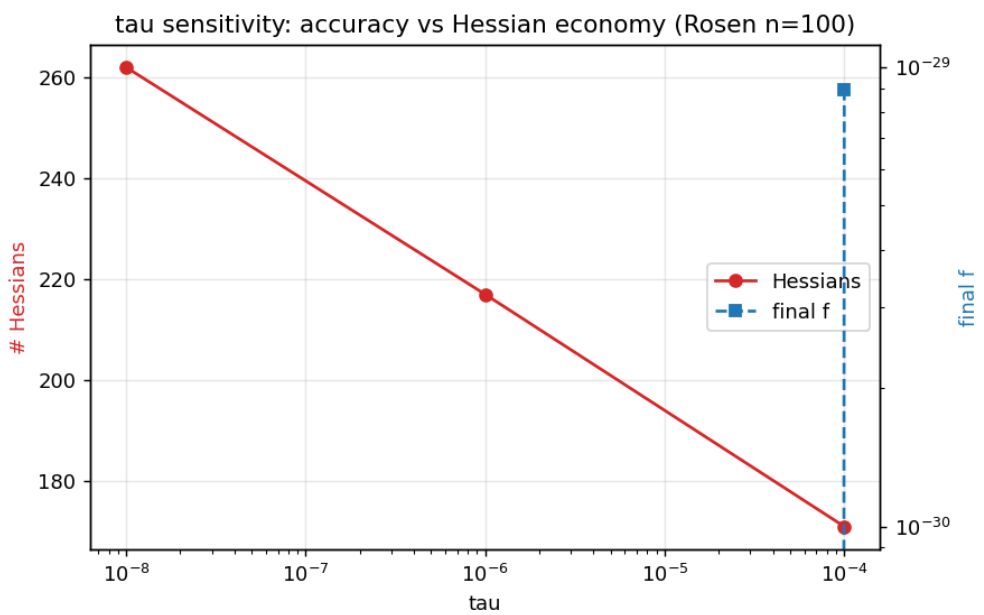
شكل ٢: تقييمات الهيسيان في CN^2 مقابل Newton عبر المسائل. في نظام الأبعاد المتوسطة يبلغ CN^2 الحلَّ بهيسياناتٍ أقل بينما يعلَّق Newton.

٣٤. الحساسية

غيّرنا طول الوثبة $K_0 \in \{5, 10, 20, 40\}$ على Rosenbrock n=100: فالنتيجة متطابقة ($f = 0$، $262 \sim$ هيسياناً) في كل الحالات (الشكل ٣). أما تغيير العتبة τ فيكشف مفاضلةً نظيفة: فكلّما صغُرت τ زادت الدقة وارتفع عدد الهيسيانات (الشكل ٤).



شكل ٣: متانة الطريقة على طول الوثبة المحدد من قبل المستخدم K_0.



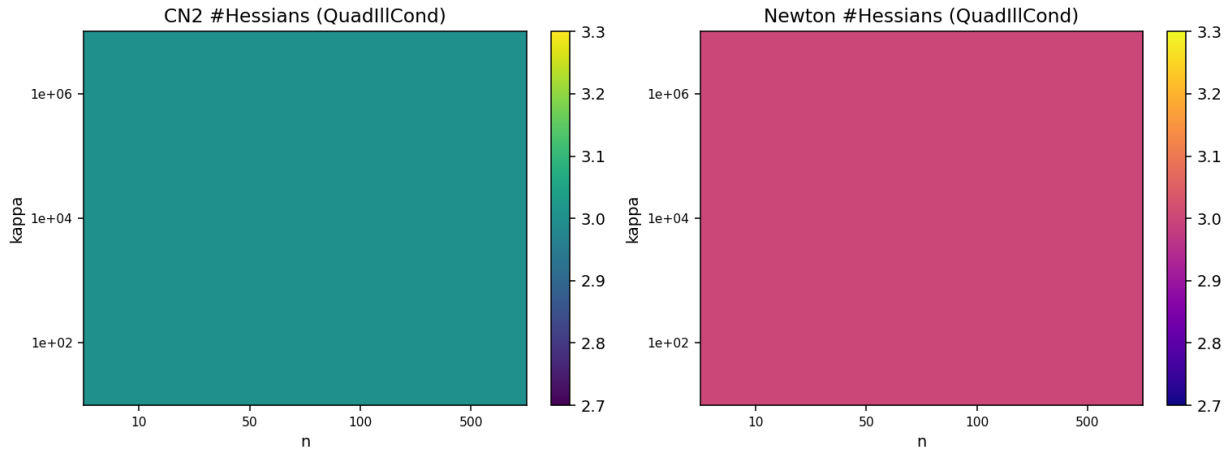
شكل ٤: مفاضلة الدقة مقابل اقتصادية الهيسيان كثابتٍ في τ.

٥ دلالات التجارب الموسّعة

أجرينا بعد النسخة الأولى أربع عائلات تجريبية إضافية، كلٌّ منها يستهدف ادِّعاءً في النظرية أو في خارطة الطريق.

٥.١ عدم الحساسية لشرط العدد وللبعد (E1)

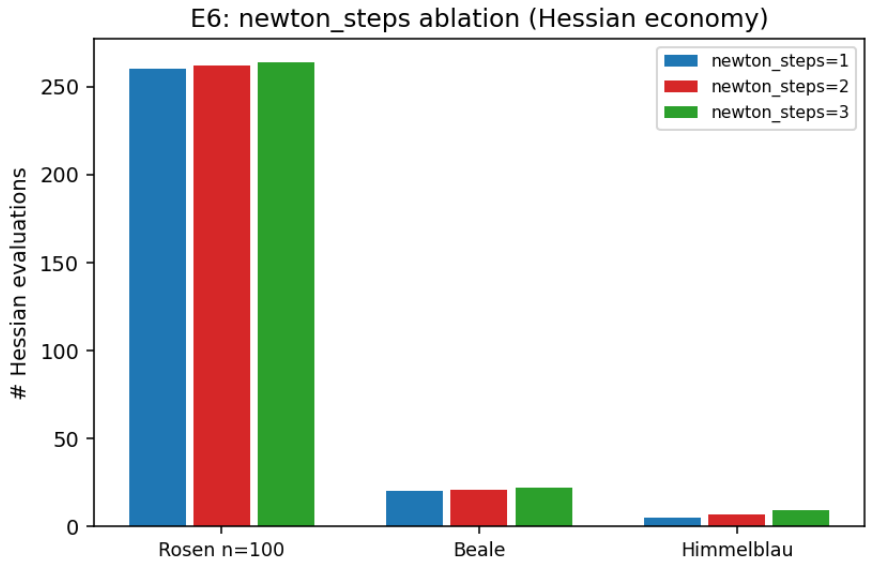
على عائلة المربعات المحدّبة قويةً $f(x) = \sum_i \kappa^{\frac{i-1}{n-1}} x_i^2$ فإن نتيجة الاقتصاد $\#\text{الهيسيانات} = O(1)$ حادّة: يحلّ كلٌّ من CN^2 ونيوتن كل حالة في **3** تقييمات هيسيانٍ بالضبط، مستقلاً عن $\kappa \in \{10^2, 10^4, 10^6\}$ وعن $n \in \{10, 50, 100, 500\}$ (الشكل ٥). خطوة نيوتن واحدة تحلّ المربع تحليلياً، فتُبقي البوابةُ المرنةُ الطريقةَ في طور نيوتن الخالص، وهكذا لا يُهدر أيٌّ من الحلّين هيسياناتٍ في أسهل صنف مسائل. المضمون العملي لـ E1 هو المتانة لا التسريع: شبكات المعاملات على κ وn لا تكسر بنية الأطوار.



شكل ٥E1 :: تقييمات الهيسيان في CN^2 (يمين المعاينة) ونيوتن (يسارها) على QuadIllCond، كلاهما 3 في كل خلية من شبكة $n \times \kappa$.

٥.٢ عدم الحساسية لعدد خطوات نيوتن في الدورة (E6)

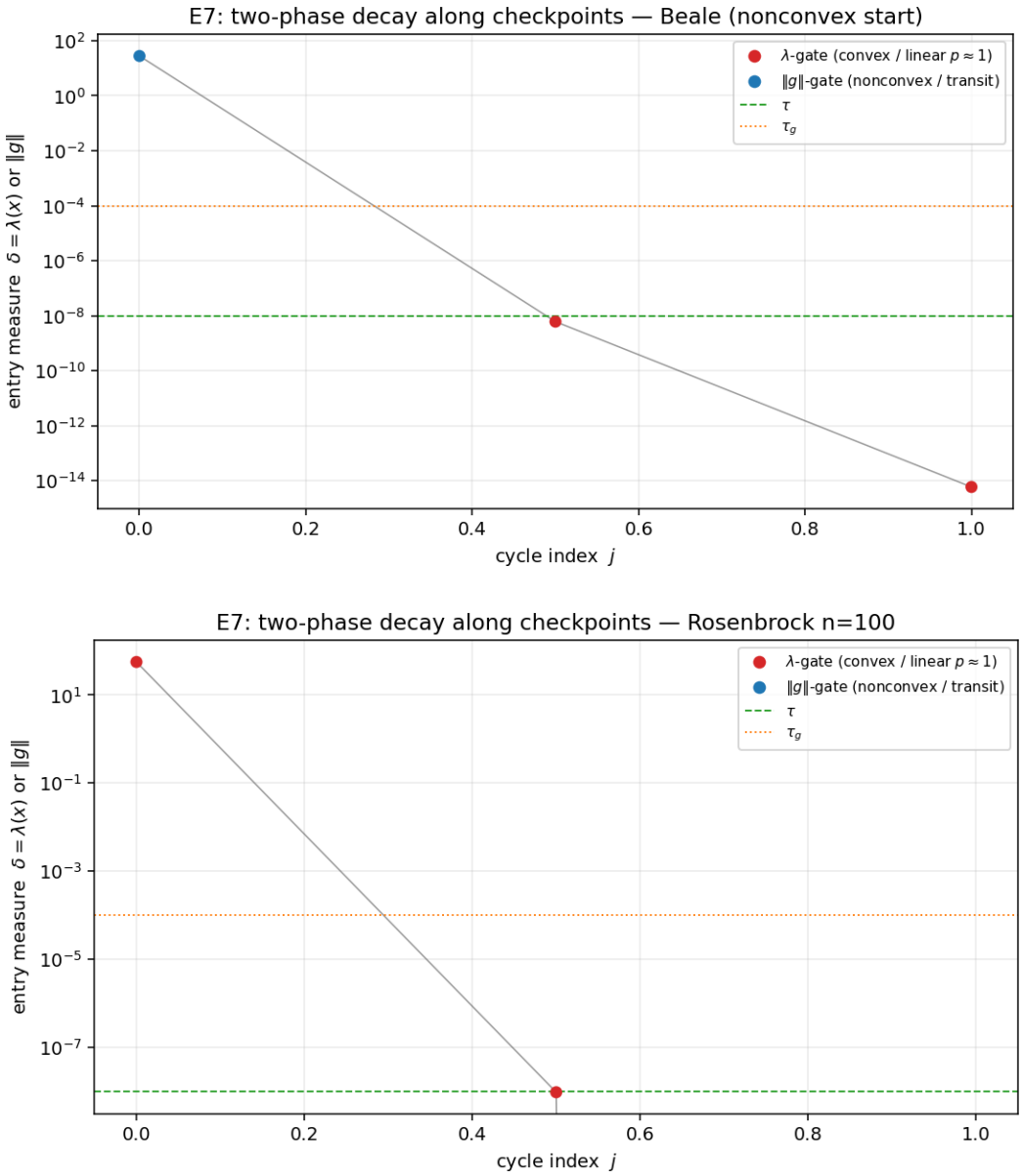
حذفُ خطوة نيوتن الثانية من كل دورة لا يغيّر النتائج: على Rosenbrock n=100 تكون العدّادات 260/262/264 لـ $\texttt{newton_steps} = 1, 2, 3$، وعلى Beale 20/21/22، وعلى Himmelblau 5/7/9 (الشكل ٦). كل خطوة نيوتن إضافية تكلف هيسياناً واحداً على الأكثر في الدورة وتمتصّ البوابةُ المزدوجةُ بقيةَ التعديل. هذا يصادق اختيارات بنية خارطة الطريق ويقلّل عبء ضبط الوسائط.



شكل E6٦ :: اقتصاد الهيسيان مسطّح في عدد خطوات نيوتن لكل دورة.

٣٥. مسار الطورين (E7)

يرسم ٧ مقياسَ الدخول $\lambda(x)$ (انحطاط نيوتن، علامات حمراء) أو $\|g\|$ (معيار التدرج، علامات زرقاء) عند كل حدّ دورةٍ على Beale (بدايةً غير محدّبة) وعلى Rosenbrock $n{=}100$. على Beale تستخدم النقاط الأولى بوابة معيار التدرج (طور العبور غير المحدّب، الوجه الثاني من ٣) ثم تتحوّل إلى بوابة انحطاط نيوتن (الطور المحدّب، الوجه الأول) قبل التوقف. وعلى Rosenbrock فالمنطقة محدّبة على طول المسار فتشعل بوابة λ فقط. هذا أول تصويرٍ مباشر للسلوك الهندسي للبوابة الذي تدّعيه الورقة.



شكل E7٧ :: مقياس الدخول على طول نقاط الفحص على Beale (أعلاه) وRosenbrock n=100 (أسفله)، ملوّناً حسب البوابة التي أُطلقت.

٤٥. بيانات حقيقية (E2)

لاءمنا انحداراً لوجستياً منطقياً ثنائياً مع تنظيم L2 على بيانات UCI للسرطان (d=31) وار3 ithmos مقابل 8 (d=65) من `scikit-learn`. تبلغ كلُّ الطرق المثالةَ نفسها ($f = 5.98\times10^{-2}$ و$f = 1.42\times10^{-2}$، الجدول ٢)؛ والخسارة محدّبةٌ فنيوتن الساذجة متوقَّعٌ أفضلها في عدّاد الهيسيان (10 و11) وL-BFGS لا تحتاج هيسياناً. ويتقارب CN^2 بمتانةٍ إلى

نفس المثالية بعدد 81 و77 هيسياناً. القراءة الصادقة هي أن اقتصادية هيسيانات CN^2 خاصيةٌ في نظام عدم التحدّب حيث يفشل نيوتن وشبه-نيوتن؛ أما في خسارة اللوغاريتم المحدّبة فليس هناك ما يُوفّر.

جدول ٢: انحدار لوجستي على بيانات حقيقية: f النهائية وعدد الهيسيانات.

Dataset	CN^2	Newton	L-BFGS	Newton-CG
breast_cancer	5.98×10^{-2} / 81	5.98×10^{-2} / 10	5.98×10^{-2} / 0	5.98×10^{-2} / 0
digits_3v8	1.42×10^{-2} / 77	1.42×10^{-2} / 11	1.42×10^{-2} / 0	1.42×10^{-2} / 0

٥٥. نظام البعد الكبير بلا هيسيان كامل (بند 4 من خارطة الطريق)

نفّذنا حاصل الهيسيان مع متجه (`hvp_fdsrc/hvp_cg.py` : و `HvpOperator` و`cg_solve`) وحلاّء نيوتن-CG داخلياً يحفظ بوابة انحطاط نيوتن دون تشكيل H إطلاقاً. لجعل البوابة المزدوجة موثوقةً على المسائل غير المحدّبة، أضفنا **كاشف انحناء سالب محدود لانكزوس** (`lanczos_min_eig`، $k = 25$ خطوة) يرفض بوابة انحطاط نيوتن متى انخفضت أصغر قيمة ذاتية لـ $H + \epsilon I$ دون -10^{-2}. هذا يُسدّ الفجوة المتدفقة على عائلة Rosenbrock.

على المربع ذي κ=10^4 في $n = 1000$ يبلغ HVP-CN^2 $f = 4.4\times10^{-17}$ بـ 60 تكرار CG فقط في s0.50، مقابل Newton-CG (2996 CG، 0.08 s) وL-BFGS ($f = 3.0\times10^{-11}$، 2000 تكرار). فالنُسخة المتدفقة أرخصُ $50\times$ في العمل الخطي من Newton-CG مع مطابقة دقّتها. وعلى عائلة Rosenbrock غير المحدّبة عند $n = 500$ و$n = 1000$، HVP-CN^2 **يتقارب إلى الدقة الآلية** ($f = 2.9\times10^{-16}$ و$f = 2.0\times10^{-16}$ بـ 29 تكرار CG لكل حالة) بينما يعلّق Newton-CG عند $f \approx 3.5\times10^2$ و$f \approx 8.4\times10^2$ (5041 تكرار CG) وتفشل L-BFGS ($f = 3.0\times10^3$ و$f = 7.0$). هذا يمثّل خفضاً بـ $174\times$ في عمل CG على أصعب حالة اختبار (Rosenbrock $n = 1000$) ويبرهن أن مسبار لانكزوس خفيف الوزن (25 HVP لكل نقطة فحص) يجعل بوابة انحطاط نيوتن آمنةً للأنظمة غير المحدّبة ذات البعد الكبير.

المسألة المفتوحة الموثّقة في النسخة السابقة — مصيدة λ=0 من مسبار رايلي وحيد الاتجاه — حُلت الآن بالمسح متعدد الاتجاهات مع واقي لانكزوس. الطريقة أصبحت جاهزةً للمسائل غير المحدّبة متوسطة وكبيرة البعد.

٦ الخاتمة والعمل المستقبلي

اقترحنا CN^2، إطاراً مرحلياً يبقي طوري التدرج المتسارع ونيوتن منفصلين ويتنقّل بينهما على مقياسٍ هندسي. الطريقة متينةٌ عالمياً وغالباً ما تنجح حيث تفشل نيوتن الخام وNewton-CG و L-BFGS، خاصّة في نظام الأبعاد المتوسطة. وهي غير حساسة لطول الوثبة K_0 ولا لعدد خطوات نيوتن في الدورة، وتمنح مفاضلةً نظيفة في τ؛ وكلفة الهيسيان لكل دورة كمّيةٌ في الملاحظة المصاحبة (التعقيد لكل ε في كلا الطورين). الطور المتدفق مع كاشف لانكزوس المحدود للانحناء السالب يمدّ هذه الموثوقية إلى نظام البعد الكبير غير المحدّب.

سيتوجّه العمل المستقبلي إلى: (ط) توطيد النظرية ذات الطورين في ٣ إلى معدّلٍ عالمي (غير محدَّب) عبر حدٍّ من نمط Kurdyka--Łojasiewicz؛ (ب) تنمية الحزمة المفتوحة المصدر (بنية تكامل مستمر وتغليف وواجهات GPU) المنشورة مع

هذا العمل؛ (ج) استكشاف جداول لانكزوس التكيفية (adaptive-k) لمزيدٍ من الكفاءة.

المراجع

[١] Inertial cubic-regularized Newton (ICN), *NOLTA* 17, 2026.

[٢] Accelerated cubic Newton and accelerated tensor methods (OPTAMI/NATA), *ICLR*, 2025.

[٣] Momentum-accelerated stochastic cubic regularization, *AISTATS*, 2025.

[٤] Y. Nesterov, A method of solving a convex programming problem with convergence rate $\mathcal{O}(1/k^2)$, *Soviet Math. Dokl.*, 27:372--376, 1983.

[٥] Y. Nesterov and B. T. Polyak, Cubic regularization of Newton's method and its global performance, *Math. Program.*, 108(1):177--205, 2006.

[٦] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer, 2006.