Voici les consignes pour votre premier devoir. à remettre le mardi 13 février. Les exercices théoriques doivent être réalisés individuellement alors que l'implantation peut être complétée en équipes. Séparez-vous le travail de codage pour en arriver au travail d'observation des algorithmes le plus rapidement possible.

1 Exercices théoriques

Exercice 1 *[Point faible de l'approximation cubique]* Exercice 2.5.8 page 75 des notes

Exercice 2 *[Convergence locale d'approximation cubique]* On considère le problème $\min_{x\in\mathbb{R}} f(x)$ et l'utilisation d'une approximation cubique dans un algorithme local. Considérons une approximation cubique de f, par exemple l'approximation de Taylor est

$$f(x+d) \approx c_x(d) \stackrel{\text{def}}{=} f(x) + f'(x)d + \frac{1}{2}f''(x)d^2 + \frac{1}{6}f'''(x)d^3.$$

Ceci correspond à une approximation quadratique de $g(x) \equiv f'(x)$

$$g(x+d) \approx q_x(d) \stackrel{\text{def}}{=} g(x) + g'(x)d + \frac{1}{2}g''(x)d^2.$$

On peut alors envisager utiliser un minimiseur d_x de $c_x(d)$ comme direction pour définir un algorithme itératif $x^+ \leftarrow x + d_x$.

a) Considérez un développement de Taylor d'ordre 3 de f, ou encore un développement d'ordre 2 de $g = f'$. En vous inspirant de la section 2.7 des notes, établissez la convergence cubique de cette méthode.

b) Considérez maintenant l'approximation cubique de la section 2.5.3.

   i) Analysez l'approximation cubique utilisant f_k, g_k, f_{k-1} et g_{k-1} ($g \equiv f'$). Montrez que le terme A se ramène à $\frac{1}{2}f''_k + \mathcal{O}(\delta_k^2)$ où $\delta_k \stackrel{\text{def}}{=} x_k - x_{k-1}$ et le terme $B = \frac{1}{6}f'''(x_k) + \mathcal{O}(\delta_k)$. Utilisez un développement de Taylor pour exprimer les quantités f_{k-1} et g_{k-1} autour de f_k et g_k ; rendez-vous jusqu'à l'ordre f''' ou g'' avant le reste.

   ii) Forts de ces estimations pour A et B, imitez le raisonnement de la section 2.8.1 pour établir l'ordre quadratique de cette variante.

Exercice 3 *[Convergence globale d'approximation cubique]* On peut utiliser ces deux approximations cubiques au sein d'une globalisation "bissection-cubique" ou bien région de confiance.

a) Est-ce le fait que $d = 0$ soit un minimum local de l'approximation de Taylor d'ordre trois est une condition nécessaire d'optimalité ? Justifiez.

b) Démontrez que la variante "bissection-cubique" converge vers un point qui satisfait aux conditions nécessaires peu importe laquelle des approximations cubiques est utilisée.

c) Démontrez que l'algorithme de région de confiance utilisant l'approximation de Taylor d'ordre 3 converge vers un point satisfaisant aux conditions nécessaires.

d) Qu'en est-il de l'algorithme région de confiance utilisant l'approximation cubique avec deux points ?

Exercice 4 *[Efficacité d'algorithmes]* L'ordre de convergence est un indice de l'efficacité d'un algorithme. Un autre indice provient de la quantité de calculs que l'algorithme doit effectuer. Si on néglige les calculs algorithmiques proprement dits en supposant que les calculs les plus lourds proviennent des évaluations de fonctions (f, f', f'', f'''), Ostrowski a proposé un *indice d'efficacité asymptotique* donné par la formule $I = \frac{\log(ordre)}{C(iter)}$. Par exemple, l'indice d'efficacité de la méthode de Newton (pure, sans globalisation) est $\log(2)/2$ et celui de la méthode de tangente $\log(\tau)$. La logique de cet indice est que la convergence est répartie équitablement sur toutes les évaluations nécessaire pour une itération. En fait, 2^I retrouve l'ordre équivalent si on effectue une seule évaluation.

a) Fournissez l'indice d'efficacité asymptotique des 8 variantes que nous avons sous la main (Newton, sécante, cubique 2 points et cubique-Taylor) sous deux formes de globalisation chacune.

b) On utilise fréquemment un test d'arrêt basé sur le fait que la fonction à minimiser f stagne, du genre $f(x_{k-1}) - f(x_k) \leq \epsilon$. Révisez vos indices d'efficacité pour y inclure ce test. Si on désire s'arrêter avec $|f'(x)| < \epsilon_g$, quelle serait une valeur raisonnable de ϵ_f pour s'arrêter lorsque $f(x_{k-1}) - f(x_k) \leq \epsilon_f$? En d'autres termes, à quoi s'attend-on comme réduction de f lorsque $|f'(x)| \searrow 0$? Un peu d'analyse sur l'ordre de convergence à zéro des quantités testées s'impose !

Exercice 5 *[Cancellation catrastrophique]* Fournissez un exemple où la cancellation catastrophique empêche l'algorithme de région de confiance de bien fonctionner alors que les formules améliorées assurent son bon fonctionnement.

2 Implantation d'algorithmes

Après avoir analysé les interpolations cubiques, mettons à l'épreuve les variantes que nous avons : Newton, sécante, cubique 2 points et cubique-Taylor. De plus, on peut les globaliser avec Trouve_intervalle + Bissection ou encore avec la stratégie de région de confiance. Nous espérons ainsi que les "vraies" performances sont cohérentes avec les prévisions données par les indices d'efficacité.

Samuel Goyette a effectué deux stages et produit les implantations que nous utiliserons pour la partie expérimentation. Voici le code disponible pour ce travail. Il ne vous reste que les versions utilisant le polynôme cubique de Taylor à compléter. Vous devrez aussi modifier les versions existantes pour y incorporer les tests d'arrêt basés sur les comparaisons d'évaluations de la fonction.

— code des variantes : `https://github.com/Goysa2/ScalarSolvers` ;
— problèmes tests : `https://github.com/JuliaSmoothOptimizers/OptimizationProblems.jl` .

a) Dans un premier temps, tester les 8 variantes sur des fonctions 1D. Implantez parcimonieusement les calculs de f, f', f''. La mesure de performance sera le nombre total d'appel à une de ces fonctions, `obj + grad + hess`. Sur au moins un exemple, exhibez l'ordre de convergence espéré ainsi que le nombre moyen d'évaluations de fonctions par itération. Utilisez comme seul test d'arrêt $|f'(x)| < \epsilon$, avec évidemment un test sur le nombre maximum d'itérations.

b) Incorporez maintenant un test d'arrêt basé sur les différences de fonction objectif pour valider les indices d'efficacité de la question 3b).