

Pokyny k vypracování semestrálního projektu k předmětu M5170 (podzim 2023)

Nutnou podmínkou pro získání úspěšného hodnocení z Matematického programování (M5170) je odevzdání vypracovaného semestrálního projektu do **5. února 2024** (v případě účasti u zkoušky před odevzdáním projektu bude hodnocení zapsáno až po odevzdání projektu – ovšem v případě nedodržení stanoveného termínu bude předmět hodnocen „X“ bez ohledu na výsledek u zkoušky).

Projekt bude věnován *numerickým metodám pro řešení úloh nepodmíněné minimalizace* a bude obsahovat analýzu dvou vybraných metod pro funkci jedné proměnné a dvou metod pro funkci více (alespoň dvou) proměnných. Zadání se *nesmí* shodovat s příklady uváděnými na přednášce, řešenými ve cvičení nebo ve sbírce příkladů. Pro funkce více proměnných také *není povolena* kvadratická funkce, tj. funkce tvaru $f(x) = x^\top Ax + b^\top x + c$.

Na každém projektu mohou pracovat *nejvýše dva studenti* (v případě lichého počtu studentů je možné povolit i jednu trojici, ale až po domluvě s přednášejícím). Konkrétní funkce si navrhnou studenti (nebo skupiny studentů) sami prostřednictvím aplikace Rozpisy v ISu (viz návod str. 3-8 níže). Návrh musí být *schválen přednášejícím* a navrhované funkce se *nesmí shodovat* s již odsouhlasenými tématy. Do kolonky „Název“ napište *navrhované funkce* (použijte stejnou/podobnou syntaxi jako v L^AT_EXu) a do kolonky „Oficiální zadání“ napište *zvolené metody*. Metody jsou rozdělené do čtyř skupin a z každé skupiny si musíte vybrat jednu metodu:

(A) **první skupina** (funkce jedné proměnné)

- (i) metoda prostého dělení (se sudým i lichým počtem bodů vyčíslení)
- (ii) metoda půlení intervalu

(B) **druhá skupina** (funkce jedné proměnné)

- (i) Fibonacciho metoda
- (ii) metoda zlatého řezu

(C) **třetí skupina** (funkce více proměnných)

- (i) metoda největšího spádu
- (ii) Newtonova metoda

(D) **čtvrtá skupina** (funkce více proměnných)

- (i) metoda sdružených gradientů (s nulováním i bez nulování volby β_k)
- (ii) metoda sdružených gradientů s alternativní volbou β_k (např. dle Polak–Ribière nebo Hestenes–Stiefel, viz přednášku nebo např. <https://m5170.page.link/RSPt>)

Po založení nového tématu mi *napište email* (adresu viz níže) s výzvou k jeho potvrzení (v případě skupinového tématu pošle výzvu každý student, neboť je potřeba potvrdit přihlášení každého studenta zvlášť). Téma musí být v ISu zadáno nejpozději **20. prosince 2023**.

V projektu analyzujte chování vybraných metod při minimalizaci zvolených funkcí. Uvažte různé hodnoty počtu vyčíslení (5, 10, 20, 50 a 100, případně i další), různé výchozí intervaly lokalizace minima, odlišné volby výchozí aproximace, atd. V úvodní části také věnujte pozornost samotným zvoleným funkcím, můžete např. analyzovat jejich průběh nebo vykreslit jejich graf apod.

Ukázku možného zpracování projektu naleznete na str. 9-38 níže. Projekt *není určen k ručním výpočtům* (ty lze obvykle použít nejvýše pro první dvě aproximace) — ke zpracování projektu můžete využít:

- (a) online maplety dostupné na MapleCloudu prostřednictvím odkazů uvedených na adrese <https://m5170.page.link/y6N7>, které
 - lze používat v on-line režimu přímo ve webovém prohlížeči (mohou se objevit problémy při kopírování výpočtů)
 - si můžete stáhnout a otevřít v Maple (pokud k němu máte přístup) nebo v MaplePlayeru dostupném *zdarma* na adrese <https://m5170.page.link/VG39> (odlehčená verze Maple=pouze "prohlížeč", dokumenty nelze upravovat – ale pro potřeby projektu to je zcela dostačující)
- (b) kódy pro Matlab (metody v \mathbb{R}) a Maple (metody v \mathbb{R}^n – vyžaduje přístup k plné verzi Maple) dostupné v jednom balíku na adrese <https://m5170.page.link/NLtk>
- (c) kódy pro R-ko (např. <https://www.r-project.org/> nebo <https://www.rstudio.com/>) dostupné na adrese <https://m5170.page.link/ZJCz> (tyto metody mohou dávat jiné aproximace než předchozí dvě varianty)
- (d) vlastní „program“ (v takovém případě uveďte podrobnosti v projektu)

Vypracované projekty se odevzdávají *elektronicky* (preferována je sazba pomocí LaTeXu) prostřednictvím *odevzdáárny* v ISu. Analýzu je vhodné doplnit i obrázky. Po vložení vypracovaného projektu do ISu mi opět *napište informační email* (tentokrát stačí pouze jeden email v případě skupinového tématu). Hodnocení, poznámky, připomínky apod. poté naleznete v příslušném *poznámkovém bloku* (o tom budete informováni emailem).

Pro zájemce: Nechcete-li se držet klasického zadání a zajímají-li Vás i jiné metody, je možné zadání projektu upravit „na míru“ *po domluvě s přednášejícím*. V projektu je např. možné analyzovat různé modifikace metod (ať již popsanych v literatuře nebo vlastních) nebo i zcela jiné numerické metody vyžadující i kompletní naprogramování. V takovém případě pak v projektu bude uveden obecný popis uvažované modifikace nebo metody včetně odkazu na relevantní literaturu (pokud se nejedná o vlastní návrh).

V případě jakéhokoli dotazu k projektu se ozvěte buď na přednášce/cvičení nebo prostřednictvím emailu uvedeného níže.


Brno 18. září 2023

doc. Mgr. Petr Zemánek, Ph.D.
zemanekp@math.muni.cz

Založení nového tématu projektu z M5170


Předměty → Matematické programování → **Rozpisy**


Moje předměty


 **M5170 Matematické programování** ★


vyučující
P. Zemánek
ROZVRH
Po 17. 9. až Pá 14. 12. Po 14:00–15:50 M2,01021 📅
SEMINÁŘ
M5170/01 Po 17. 9. až Pá 14. 12. Po 16:00–16:50 M2,01021 📅, P. Zemánek


5 kr. | zk


 Diskusní fórum

 Studijní materiály 2

 Odevzdávárny 1

 Poskytovna

 Interaktivní osnova

 Poznámkové bloky

• Seminární skupiny

• Zkoušky

• Rozpisy


• Katalog

• Spolužáci

Odkaz **založit**

Přehled témat

PřF: **M5170** Matematické programování (podzim 2018)

Rozpisy  **Témata:** založit, přihlašování


M5170: Projekt

Řadit témata dle: názvu | poslední modifikace | **vedoucího**

Zobrazit témata: moje aktuální | aktuální volná | **všetchna aktuální** | pokročilý výběr ▾

doc. Mgr. Petr Zemánek, Ph.D.

1. $f(x) = \exp(x) - 5 \cdot x + 5$, $f(x,y) = 4 \cdot x^2 + y^2 + \sin(x + y)$

☐ Vede: doc. Mgr. Petr Zemánek, Ph.D., učo 78442 



Vyplňte **název projektu** (zvolené funkce) a **oficiální zadání** (zvolené metody)

Rozpisy: úprava rozpisu



Témata: **založit** (hromadně), úprava témat, studenti

M5170: Projekt

Zavádění nového tématu

Název:

$f(x) = x^4 - \sqrt{\text{abs}(x-1)} + 4x + 12$; $f(x,y) = \cos(x) + y^2 + y^4$

Název anglicky:

Další názvy ▾

Oficiální zadání: ?

Budeme se věnovat funkcím:

$f(x) = x^4 - \sqrt{\text{abs}(x-1)} + 4x + 12$

$f(x,y) = \cos(x) + y^2 + y^4$

Zvolené metody:

(A) metoda půlení intervalů

(B) metoda zlatého řezu

(C) metoda největšího spádu

(D) metoda sdružených gradientů

Předběžné zadání ▾

Nastavte **vedoucího** tématu (vždy doc. Mgr. Petr Zemánek, Ph.D.; do vyhledávacího pole zadejte UČO 78442)

Oficiální zadání: ?

Budeme se věnovat funkcím:

$$f(x) = x^4 - \sqrt{\text{abs}(x-1)} + 4x + 12$$

$$f(x,y) = \cos(x) + y^2 + y^4$$

Zvolené metody:

(A) metoda půlení intervalů

(B) metoda zlatého řezu

(C) metoda největšího spádu

(D) metoda sdružených gradientů

Předběžné zadání ▾

Dohledání osob

Vedoucí, oponent, konzultant, ...

Vyhledat osoby



Vyhledejte osobu podle příjmení či uča.

Označením sloupce vlevo osobu odstraníte.

Vedoucí, oponent, konzultant, ...

☒ doc. Mgr. Petr Zemánek, Ph.D. (ÚMS Ústavy PřF MU), učo 78442

vedoucí ▾

Literatura ▾

Vše **ostatní ponechte** v původním nastavení a téma **uložte**

Vedoucí, oponent, konzultant, ...

<input checked="" type="checkbox"/>	doc. Mgr. Petr Zemánek, Ph.D. (ÚMS Ústavy PFF MU), učo 78442	vedoucí
-------------------------------------	--	---------



Literatura ▾


Poznámky ▾


Další údaje ▾

Zkratka tématu:

Související webová stránka: (zadejte celé URL včetně např. http://stroj.cz/)


K tématu se lze přihlásit od:  hod. do  vč.



K tématu se student smí přihlásit, jen když splňuje tyto předpoklady:
Zadejte prerekvizity jako pro zápis předmětu. (Návod. 

Maximum přihlášených studentů: (Nezadáte-li kapacitu, není omezeno)

☐ vytvořit studijní kruh

☐ přihlášení studenta k tématu musí schválit vedoucí 

☒ přihlášení k tématu oznámit e-mailem

☒ toto téma již zveřejnit studentům

☒ **Uložit** Uložit a vytvořit další

Pošlete email ohledně potvrzení přihlášení k tématu

doc. Mgr. Petr Zemánek, Ph.D.

1. $f(x)=\sqrt{x^4+1}-2x+3, f(x,y)=\ln(x^2+y^4+x+3)$

☐ Vede: doc. Mgr. Petr Zemánek, Ph.D., učo 78442

Studenti (max. 4):
1.
2.
3.
4.

čeká se na potvrzení přihlášení k tématu

Studenti se mohou přihlašovat do 11. 12. 2018 vč.

Zobrazit operace

Takto vypadá potvrzené přihlášení

doc. Mgr. Petr Zemánek, Ph.D.

1. $f(x)=\sqrt{x^4+1}-2x+3, f(x,y)=\ln(x^2+y^4+x+3)$

☐ Vede: doc. Mgr. Petr Zemánek, Ph.D., učo 78442

Studenti (max. 4):
1.

Klá

roč 3]

2.
3.
4.

Studenti se mohou přihlašovat do 11. 12. 2018 vč.

Zobrazit operace

Postup v případě tématu pro více studentů: Téma do rozpisu zadá jeden ze studentů (dle návodu výše) a ostatní studenti se k němu pouze přihlásí.

Projekt – numerické metódy

Matematické programovanie

Martin Dettelbacher, Silvia Hnilická, Simona Kolenčíková

Obsah

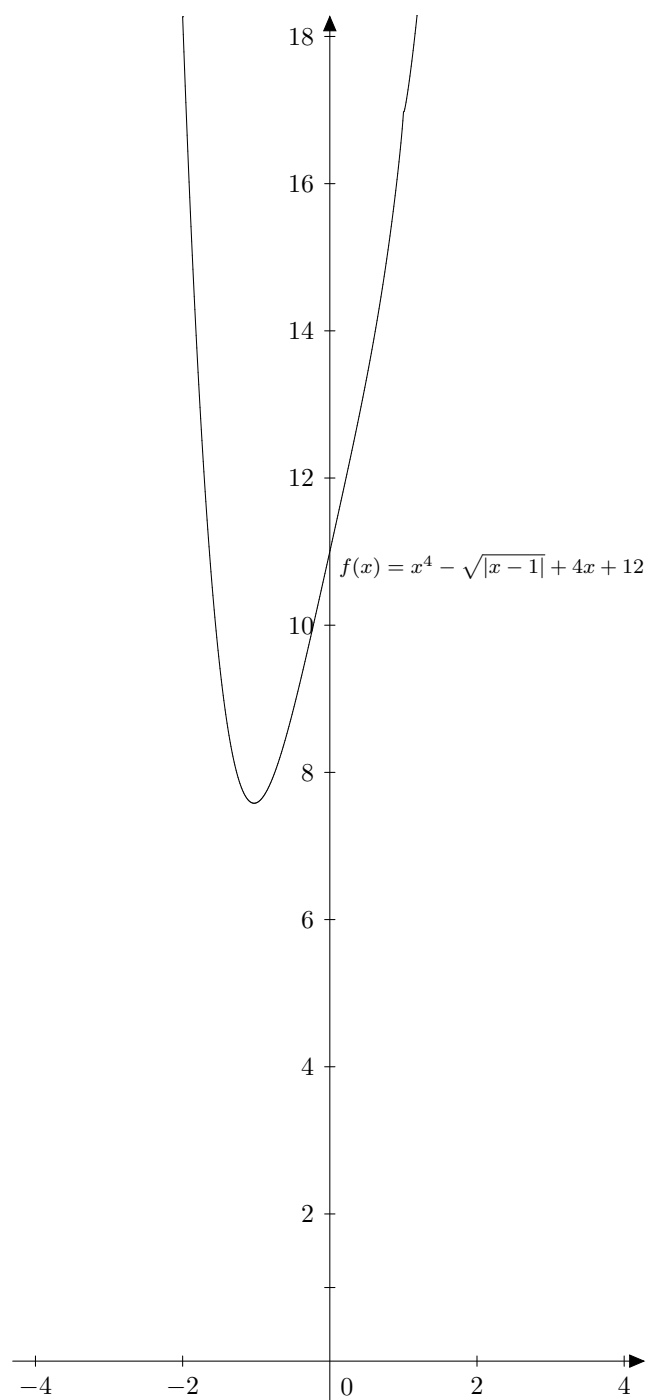
Analýza numerických metod pro funkci jedné proměnné	3
1.1 Metoda půlení intervalu	5
1.1.1 Analýza chování metody při různých hodnotách počtu vyčíslení . .	5
1.1.2 Analýza chování metody při různých výchozích intervalech lokalizace minima	7
1.2 Metoda zlatého řezu	10
1.2.1 Analýza chování metody při různých hodnotách počtu vyčíslení . .	10
1.2.2 Analýza chování metody při různých výchozích intervalech lokalizace minima	12
Analýza numerických metod pro funkci více proměnných	14
1.3 Metóda najväčšieho spádu	16
1.3.1 Teória	16
1.3.2 Analýza metódy pri rôznych hodnotách počiatočnej aproximácie . .	18
1.4 Metóda združených gradientov	24
1.4.1 Teória	24
1.4.2 Analýza chovania metódy pri rôznych hodnotách počiatočného bodu	25

Analýza numerických metod pro funkci jedné proměnné

Základní informace

K analýze chování numerických metod pro funkci jedné proměnné jsme si zvolili funkci $f(x) = x^4 - \sqrt{|x-1|} + 4x + 12$ (na následující straně je tato funkce vykreslena pro omezený rozsah hodnot x a y). „Přesného“ minima tato funkce nabývá v bodě $x^* \doteq -1.0284390691$, přičemž $f(x^*) \doteq 7.5807125417$.

Metody, které na této funkci budeme zkoumat, jsou **metoda půlení intervalu** a **metoda zlatého řezu**.



Obrázek 1: Graf funkce $f(x) = x^4 - \sqrt{|x-1|} + 4x + 12$

1.1 Metoda půlení intervalu

1.1.1 Analýza chování metody při různých hodnotách počtu vyčíslení

Zvolíme si pevný výchozí interval lokalizace minima (ILM) a budeme sledovat, jak se liší dosažená přesnost a „efektivita“ (dosažená přesnost vzhledem k počtu vyčíslení) při různých hodnotách počtu vyčíslení.

Vzhledem k tomu, že počet vyčíslení u metody půlení intervalu musí být sudý, vyzkoušíme postupně hodnoty počtu vyčíslení 4, 10, 20, 50 a 100.

Pokud jde o volbu (pevného) výchozího ILM, byli bychom rádi, kdyby v tuto chvíli nebyl v žádném smyslu „patologický“. Nabízí se proto vzít tento interval zhruba symetrický kolem bodu skutečného minima $x^* \doteq -1.0284390691$. Současně požadujeme, aby výchozí ILM byl dostatečně dlouhý – při krátkém ILM by zanikaly rozdíly mezi jednotlivými počty vyčíslení. Jako výchozí ILM tedy zvolme např. interval $I_0 := [-10, 9]$.

Konečně zbývá zvolit vhodné $\delta \in (0, \frac{1}{2}(b-a)) = (0, \frac{19}{2})$. Je vidět (např. ze vztahu pro délku ILM po $N = 2k$ vyčísleních $l_N = \frac{b-a}{2^k} + \frac{2^k-1}{2^{k-1}}\delta$), že s větším δ se zvětšují délky ILM. Je tudíž efektivní volit δ dostatečně malé. Zvolme např. $\delta := 0.02$.

Nyní se již budeme věnovat minimalizaci funkce f metodou půlení intervalu pro různé počty vyčíslení $N \in \{4, 10, 20, 50, 100\}$.

N = 4

Po $N = 4$ vyčísleních dostaneme ILM $I_2 = [a_2, b_2] = [\lambda_2, \mu_1] = [-5.26, -0.48]$, tudíž apriorní odhad maximální chyby je $\varepsilon_0 = \frac{1}{2}l_2 = 2.39$. Aproximace bodu minima je $\bar{x} = \frac{1}{2}(a_2 + b_2) = -2.87$ a skutečného minima f nabývá v $x^* \doteq -1.0284390691$, takže skutečná chyba je $|\bar{x} - x^*| \doteq 1.841561$.

Skutečná chyba se vzhledem k velmi nízkému počtu vyčíslení a k relativně širokému výchozímu ILM nemusí zdát příliš velká, avšak kdyby nás zajímala hodnota funkce f v bodě svého minima, tak se aproximací pomocí $\bar{x} = -2.87$ dopustíme již dost velké nepřesnosti, neboť $f(\bar{x}) \doteq 66.39929$, zatímco $f(x^*) \doteq 7.580712541$.

N = 10

Po $N = 10$ vyčísleních získáme ILM $I_5 = [-1.1125, -0.48]$, apriorní odhad chyby $\varepsilon_0 = 0.31625$, aproximaci bodu minima $\bar{x} = -0.79625$ a skutečnou chybu $|\bar{x} - x^*| \doteq 0.232189$. Funkční hodnota v nalezené aproximaci je $f(\bar{x}) \doteq 7.87673$.

Přidáním šesti dalších bodů vyčíslení se tedy zmenšila skutečná chyba $|\bar{x} - x^*|$ o 1.609372 (z původní hodnoty 1.841561 na 0.232189) a funkční hodnota $f(\bar{x}) \doteq 7.87673$ se výrazně přiblížila funkční hodnotě ve skutečném minimu $f(x^*) \doteq 7.580712541$.

To, že se přidáním pouhých šesti vyčíslení hodnota $f(\bar{x})$ hodně zpřesnila, je dáno prudkým sklonem funkce f a volbou výchozího ILM $I_0 = [-10, 9]$. Střed tohoto intervalu, bod

$s_0 = -0.5$, a také oba body $\lambda_1 = s_0 - \delta = -0.5 - 0.02 = -0.52$ a $\mu_1 = s_0 + \delta = -0.5 + 0.02 = -0.48$ leží v pravém okolí bodu minima $x^* \doteq -1.0284390691$, kde je funkce f rostoucí, tj. $f(\lambda_1) < f(\mu_1)$, a proto se hned v prvním kroku interval $I_0 = [-10, 9]$ zkrátí zprava na $I_1 = [-10, \mu_1] = [-10, -0.48]$. Střed tohoto intervalu je hodně vychýlený od bodu skutečného minima $x^* = -1.0284390691$, takže se následujícími vyčísleními přibližujeme k bodu x^* pouze zleva, kde funkce f strmě klesá.

V dalším lze očekávat, že ani přidáním velkého počtu vyčíslení se aproximace \bar{x} ani funkční hodnota $f(\bar{x})$ dramaticky měnit nebudou.

N = 20

Po $N = 20$ vyčísleních obdržíme ILM $I_{10} = [-1.056953125, -0.9984375]$, apriorní odhad chyby $\varepsilon_0 = 0.0292578125$, aproximaci bodu minima $\bar{x} = -1.0276953125$ a skutečnou chybu $|\bar{x} - x^*| \doteq 0.000743756583277966$. Funkční hodnota v nalezené aproximaci je $f(\bar{x}) \doteq 7.58071607448856$.

Přidáním deseti dalších bodů vyčíslení se skutečná chyba $|\bar{x} - x^*|$ zmenšila už jen asi o 0.231445243 (z předchozí hodnoty 0.232189 na 0.000743756583277966) a funkční hodnota $f(\bar{x}) \doteq 7.58071607448856$ již aproximuje skutečnou hodnotu minima $f(x^*) \doteq 7.580712541$ s přesností řádově 10^{-5} .

Zatímco v prvních pěti krocích (deseti vyčísleních) se metoda „vypořádávala“ s velkým vychýlením nalevo od bodu x^* , způsobeným hned prvním zkrácením výchozího ILM $I_0 = [-10, 9]$ zprava, tak v šestém kroku se už vyčíslení $\lambda_6 = -0.81625$, $\mu_6 = -0.77625$ dostala napravo od $x^* = -1.0284390691$. V následujících krocích se díky tomu příslušné ILM už zkracovaly z obou stran, což uspokojivě vyrušilo vlivy počátečního vychýlení a poskytlo malou skutečnou chybu. Při zvoleném I_0 tak bude počet vyčíslení $N = 20$ patrně (z námi uvažovaných počtů vyčíslení) ideálním kompromisem mezi přesností a výpočetní náročností.

N = 50

Po $N = 20$ vyčísleních je ILM $I_{25} = [-1.04831065416336, -1.00831008911133]$, apriorní odhad chyby $\varepsilon_0 = 0.0200002825260162$, aproximace bodu minima $\bar{x} = -1.02831037163734$ a skutečná chyba $|\bar{x} - x^*| \doteq 0.000128697445933357$. Funkční hodnota v nalezené aproximaci je $f(\bar{x}) \doteq 7.58071264755121$.

Ačkoli jsme přidali hned 30 bodů vyčíslení, skutečná chyba $|\bar{x} - x^*|$ je nyní nižší zhruba o zanedbatelných $6 \cdot 10^{-4}$ (předchozí hodnota, tj. pro $N = 20$, byla 0.000743756583277966 , současná je 0.000128697445933357) a hodnota $f(\bar{x}) \doteq 7.58071264755121$ se od té předchozí liší jen v řádu 10^{-6} . Od skutečné minimální hodnoty $f(x^*) \doteq 7.580712541$ se nyní $f(\bar{x})$ liší až v řádu 10^{-7} .

Chyba při aproximaci funkční hodnoty $f(x^*)$ číslem $f(\bar{x})$ je nižší než chyba při aproximaci x^* číslem \bar{x} , protože \bar{x} leží ve velmi blízkém okolí (stacionárního) bodu x^* a na tomto okolí $f(x)$ roste (resp. klesá) pomaleji než funkce $g(x) = x$ (resp. $g(x) = -x$).

N = 100

Výsledkem $N = 100$ vyčíslení je ILM $I_{50} = [-1.04831063492939, -1.00831063492938]$, apriorní odhad chyby $\varepsilon_0 = 0.02000000000000084$, aproximace bodu minima $\bar{x} = -1.02831063492938$ a skutečná chyba $|\bar{x} - x^*| \doteq 0.000128434153894519$. Funkční hodnota v nalezené aproximaci je $f(\bar{x}) \doteq 7.58071264711869$.

V souladu s našimi očekáváními došlo i přes zdvojnásobení už tak vysokého počtu vyčíslení jen k minimálnímu zpřesnění výpočtů, shrňme tedy jen to nejdůležitější:

	aproximace \bar{x}	skutečné minimum x^*
bod minima	-1.02831063492938	-1.0284390691
hodnota f v bodě minima	7.58071264711869	7.5807125417

Možná stojí za povšimnutí také hodnota apriorního odhadu chyby $\varepsilon_0 = 0.02000000000000084$ – jedná se skoro přesně o hodnotu zvoleného $\delta = 0.02$. Jednotlivých zúžení ILM proběhlo tolik, že v 50. kroku získáme ILM, jenž má na každou stranu od svého středu délku téměř přesně δ , tj. celkovou délku má skoro přesně 2δ . To současně znamená, že po nějakém počtu kroků $k > 50$ už se při daném $\delta = 0.02$ aproximace \bar{x} nebude moci dále zlepšovat. Konkrétně tímto k je shodou okolností hned $k = 51$, tj. od 101. vyčíslení dále už se hodnoty $\lambda_i, \mu_i, i = 51, 52, \dots$, a krajní body příslušných ILM neustále opakují. K nejpřesnější aproximaci, které pro zvolené I_0 a δ lze dosáhnout, tedy potřebujeme $N = 102$ vyčíslení a hodnota této aproximace je $\bar{x}_{102} = -1.02831063492938$.

1.1.2 Analýza chování metody při různých výchozích intervalech lokalizace minima

Dále budeme různě měnit výchozí ILM a budeme sledovat, jaký mají tyto změny vliv na práci metody půlení intervalu.

Přesnost metody bude (kromě počtu vyčíslení) zřejmě ovlivněna jak délkou výchozího ILM, tak i jeho polohou (vzhledem k bodu skutečného minima x^*). Zaměříme se na oba tyto vlivy.

Dopady změn výchozího ILM budeme zkoumat pro počty vyčíslení $N \in \{4, 10, 20, 50, 100\}$, s nimiž jsme už pracovali v předchozím odstavci (abychom měli srovnání). Hodnotu $\delta = 0.02$ ponecháme.

Změny délky výchozího ILM

První, co se nabízí, je zkrátit výchozí ILM $I_0 = [-10, 9]$ z předchozího odstavce např. na polovinu (beze změny relativní polohy). Délka I_0 je 19, tudíž z obou stran interval I_0 zkrátíme o $\frac{19}{4} = 4.75$. Získáme tak $I_0^1 = [-5.25, 4.25]$.

Dosadíme-li do vztahu pro délku ILM po $N = 2k$ vyčísleních ($l_N = \frac{b-a}{2^k} + \frac{2^k-1}{2^{k-1}}\delta$) za $b-a$ výraz $\frac{1}{2}(b-a)$, dostaneme pro délku ILM po N vyčísleních s polovičním výchozím ILM

$l_N^1 = \frac{b-a}{2^{k+1}} + \frac{2^k-1}{2^{k-1}}\delta$. Z tohoto vztahu je vidět, že pro malé počty kroků k (resp. malé počty vyčíslení N) přinese poloviční výchozí ILM podstatné zpřesnění – díky výrazně kratšímu ILM po N vyčísleních jsou pravděpodobné přesnější aproximace \bar{x} . Nicméně pro velká k je (v případě l_N i l_N^1) délka ILM po $N = 2k$ vyčísleních určena spíše druhým sčítancem (který je v obou případech stejný); první sčítanec pro velká k jde k nule.

Očekáváme proto, že pro malá N bude výchozí ILM $I_0^1 = [-5.25, 4.25]$ poskytovat výrazné zpřesnění při hledání bodu minima x^* , pro velká N by rozdíly mezi použitím I_0 a I_0^1 měly zanikat. Zda tomu tak ve skutečnosti je, ukazují následující tabulky (pro srovnání připomínáme, že $x^* = -1.0284390691$).

N = 4	\bar{x}	$ \bar{x} - x^* $
$I_0 = [-10, 9]$	-2.87	1.84156093091672
$I_0^1 = [-5.25, 4.25]$	-1.6825	0.654060927394754
N = 10	\bar{x}	$ \bar{x} - x^* $
$I_0 = [-10, 9]$	-0.79625	0.232189069083278
$I_0^1 = [-5.25, 4.25]$	-0.9434375	0.0850015726052457
N = 20	\bar{x}	$ \bar{x} - x^* $
$I_0 = [-10, 9]$	-1.0276953125	0.000743756583277966
$I_0^1 = [-5.25, 4.25]$	-1.031201171875	0.00276209926975413
N = 50	\bar{x}	$ \bar{x} - x^* $
$I_0 = [-10, 9]$	-1.02831037163734	0.000128697445933357
$I_0^1 = [-5.25, 4.25]$	-1.02831068485975	0.000128387745493397
N = 100	\bar{x}	$ \bar{x} - x^* $
$I_0 = [-10, 9]$	-1.02831063492938	0.000128434153894519
$I_0^1 = [-5.25, 4.25]$	-1.02831063492938	0.000128437675866788

Výsledky tedy jsou v souladu s tím, co jsme čekali. V některých případech (např. $N = 20$) je dokonce vyšší přesnosti dosaženo při dvakrát širším výchozím ILM I_0 , což názorně potvrzuje, že přesnost při konkrétním počtu vyčíslení není předurčena pouze šířkou výchozího ILM.

Dělení délky výchozího ILM I_0 jinými čísly by vedlo na podobnou analýzu a zkracování I_0 z různých stran o různé hodnoty vede na otázku relativní polohy výchozího ILM vzhledem k bodu minima.

Změny polohy výchozího ILM vzhledem k bodu minima x^*

V předchozím odstavci jsme narazili na problém, že se střed intervalu $I_0 = [-10, 9]$ trefil blízko bodu skutečného minima x^* , což bylo příčinou velkého vychýlení středu následujícího ILM od bodu x^* a toto vychýlení (při malém počtu vyčíslení N) vedlo k nepřesné aproximaci \bar{x} .

Přirozeně tak vzniká otázka, jaká poloha na číselné ose výchozího ILM dané délky při daném počtu vyčíslení a daném $\delta = 0.02$ je u metody půlení intervalu výhodná.

Známe-li bod skutečného minima x^* (což v praxi příliš nedává smysl), lze stanovit „ideální“ polohu výchozího ILM (při jeho dané délce, daném $N = 2k$ a δ) „zpětně“.

V takovém případě totiž víme, že poslední ILM musí být tvaru $[x^* - \frac{l_N}{2}, x^* + \frac{l_N}{2}]$. Můžeme si stanovit, že např. levý krajní bod tohoto intervalu bude zároveň levým krajním bodem výchozího ILM, tj. že se po celou dobu jednotlivé ILM budou zkracovat zprava. Pak už si stačí jen uvědomit, že pohybujeme-li se napravo od x^* , platí vždy $f(\lambda_k) < f(\mu_k)$, a tak dojdeme k algoritmu, který bychom mohli popsat zhruba takto:

- $s_k := x^*$ (s_k je střed ILM v k -tém, tj. posledním kroku)
- $ILM_k = [\overbrace{s_k - \frac{l_N}{2}}{=:a}, s_k + \frac{l_N}{2}]$
- $s_{k-1} = (s_k + \frac{l_N}{2}) - \delta$
- $ILM_{k-1} = [a, s_{k-1} + (s_{k-1} - a)] = [a, 2s_{k-1} - a]$
- $s_{k-2} = s_{k-1} + (s_{k-1} - a) - \delta = 2s_{k-1} - a - \delta$
- $ILM_{k-2} = [a, s_{k-2} + (s_{k-2} - a)] = [a, 2s_{k-2} - a] = [a, 2(2s_{k-1} - a - \delta) - a] = [a, 4s_{k-1} - 3a - 2\delta]$ atd.

Takto dojdeme k rekurentním vztahům ($m \in \mathbb{Z}$, $0 \leq m \leq k$):

$$s_{k-m} = 2^{m-1}s_{k-1} - (2^{m-1} - 1)a - (2^{m-1} - 1)\delta = 2^{m-1}[(s_k + \frac{l_N}{2}) - \delta] - (2^{m-1} - 1)a - (2^{m-1} - 1)\delta$$

$$ILM_{k-m} = [a, 2^m s_{k-1} - (2^m - 1)a - (2^m - 2)\delta] = [a, 2^m [(s_k + \frac{l_N}{2}) - \delta] - (2^m - 1)a - (2^m - 2)\delta].$$

Vrátíme-li se ale k otázce, jak nejlépe umístit např. výchozí ILM délky 19 pro $N = 4$ vyčíslení a $\delta = 0.02$, známe-li $x^* = 1.028490691$, tak uvedené vztahy samozřejmě používat nemusíme a stačí položit $a := s_2 - \frac{l_4}{2} = -3.418490691$ a $b := a + 19$, dostaneme $ILM_0 = [-3.418490691, 15.58150931]$.

Už bylo naznačeno, že provedená analýza v praxi nejspíš není příliš užitečná, protože bod x^* bývá právě tím, co neznáme (a chceme odhadnout). V naší situaci bod x^* ovšem známe, takže nám krátké zamyšlení nad zpětným hledáním ideálního výchozího ILM (včetně jeho následných ILM) přišlo zajímavější než náhodné zkoušení.

1.2 Metoda zlatého řezu

1.2.1 Analýza chování metody při různých hodnotách počtu vyčíslení

V případě metody zlatého řezu analýzu pojmem o něco stručněji a využijeme toho, že již máme nějaké výsledky z analýzy metody půlení intervalu. Zaměříme se proto spíše na srovnávání metody zlatého řezu a metody půlení intervalu.

Z tohoto důvodu budeme opět uvažovat hodnoty počtu vyčíslení $N \in \{4, 10, 20, 50, 100\}$, výchozí ILM $I_0 = [-10, 9]$ a (v případě metody půlení intervalu) $\delta = 0.02$.

Znovu se krátce zamysleme nad tím, co můžeme očekávat. U metody zlatého řezu je délka ILM po k -tém kroku (tj. po $k+1$ vyčísleních) $l_k = \frac{l_0}{\tau^k}$, kde l_0 je délka výchozího ILM I_0 .

Pro rychlost konvergence zde tedy máme: $\lim_{k \rightarrow \infty} \frac{l_{k+1}}{l_k} = \lim_{k \rightarrow \infty} \frac{\frac{l_0}{\tau^{k+1}}}{\frac{l_0}{\tau^k}} = \lim_{k \rightarrow \infty} \frac{1}{\tau} = \frac{1}{\tau} \approx$

0.618. U metody půlení intervalu takto univerzálně rychlost konvergence určit nemůžeme, neboť pro $k \rightarrow \infty$ délky ILM konvergují k číslu δ a uvedená limita by vyšla 1. Pro počty vyčíslení, které uvažujeme, však velmi snadno nahlédneme, že každý následující ILM je oproti předchozímu (při dostatečně malém δ) zhruba poloviční, takže rychlost konvergence je ≈ 0.5 . Mohlo by se tak zdát, že hledání minima metodou půlení intervalu je rychlejší než hledání minima metodou zlatého řezu, ovšem musíme vzít v potaz, že zatímco metodou půlení intervalu pro $N = 2k$ vyčíslení obdržíme právě $k = \frac{N}{2}$ ILM, metoda zlatého řezu poskytne (kromě prvního kroku) s každým novým vyčíslením nový ILM, tudíž pro N vyčíslení získáme hned $N - 1$ ILM. Krom toho u metody zlatého řezu není maximální dosažitelná přesnost limitována číslem δ .

Díky tomu očekáváme, že výsledky dosažené metodou zlatého řezu budou přesnější. Srovnání výsledků metody zlatého řezu s (již získanými) výsledky metody půlení intervalu uvádíme v následujících tabulkách (připomínáme, že $I_0 = [-10, 9]$, $x^* = 1.028490691$, $f(x^*) = 7.5807125417$ a u MPI je $\delta = 0.02$).

N = 4	MZŘ	MPI
\bar{x}	-0.4999999999999999	-2.87
$ \bar{x} - x^* $	0.528439069083279	1.84156093091672
$f(\bar{x})$	8.83775512860841	66.3992900527094
poslední ILM	$[-2.742645786248, 1.742645786248]$	$[-5.26, -0.48]$

Vidíme, že i pro nízké $N = 4$, kdy metodou zlatého řezu získáme oproti MPI jen o jeden ILM navíc, už MZŘ poskytuje znatelně přesnější výsledky.

Při takto nízkých počtech vyčíslení jsou ještě (absolutní) délky, o které se ILM zkracují, dost velké, takže i jedno zkrácení navíc znamená podstatné zpřesnění a uvedené srovnání tak není příliš překvapivé.

N = 10	MZŘ	MPI
\bar{x}	-0.999913464864143	-0.79625
$ \bar{x} - x^* $	0.0285256042191355	0.232189069083278
$f(\bar{x})$	7.5807125417	7.87673131607973
poslední ILM	$[-1.12489183108018, -0.874935098648107]$	$[-1.1125, -0.48]$

Pro $N = 10$ vyčíslení nám MZŘ dá už o čtyři ILM více než MPI a přesnější aproximace pomocí MZŘ je opět zcela v souladu s očekáváním.

N = 20	MZŘ	MPI
\bar{x}	-1.02840070355957	-1.0276953125
$ \bar{x} - x^* $	$3.83655237110592e - 05$	0.000743756583277966
$f(\bar{x})$	7.58071255113656	7.58071607448856
poslední ILM	$[-1.02941685500799, -1.02738455211114]$	$[-1.056953125, -0.9984375]$

Při počtu vyčíslení $N = 20$ sice MPI zaznamenala velký skok v přesnosti oproti předchozímu $N = 10$, ale přesnost dosažená pomocí MZŘ je opět vyšší.

Dokonce už při tomto počtu vyčíslení je metodou zlatého řezu dosaženo lepší přesnosti, než jaká je metodou půlení intervalu vůbec dosažitelná (při daném δ a I_0).

N = 50	MZŘ	MPI
\bar{x}	-1.02843906363071	-1.02831037163734
$ \bar{x} - x^* $	$5.45256639838954e - 09$	0.000128697445933357
$f(\bar{x})$	7.58071254173211	7.58071264755121
poslední ILM	$[-1.02843906417688, -1.02843906308454]$	$[-1.04831065416336, -1.00831008911133]$

V důsledku vysokého počtu vyčíslení $N = 50$ už u MPI nedochází k velkým zpřesněním, poněvadž poslední ILM má nyní délku téměř 2δ a jednotlivé ILM už „se nemají kam zkracovat“. Naproti tomu u MZŘ skutečná chyba aproximace klesla řádově na $5 \cdot 10^{-9}$, takže už je naprosto zanedbatelná.

N = 100	MZŘ	MPI
\bar{x}	-1.0284390632925	-1.02831063492938
$ \bar{x} - x^* $	$5.79078052176385e - 09$	0.000128434153894519
$f(\bar{x})$	7.58071254173211	7.58071264711869
poslední ILM	$[-1.0284390632925, -1.0284390632925]$	$[-1.04831063492939, -1.00831063492938]$

Pro $N = 100$ už MPI v podstatě „narazila na své limity přesnosti“, jak jsme diskutovali dříve, a MZŘ „narazila na limity“ výpočetní přesnosti systému MATLAB, takže se výsledky v obou případech skoro nezměnily.

Počínaje 75. ILM totiž u MZŘ splynuly levý a pravý krajní bod posledního ILM a dalšího zpřesnění tak nelze dosáhnout (bez případných úprav nastavení systému MATLAB, do kterých jsme se nechtěli pouštět).

1.2.2 Analýza chování metody při různých výchozích intervalech lokalizace minima

Na závěr kapitoly pro odlehčení vyzkoušíme, jaká aproximace \bar{x} bude dosažena pro výchozí ILM, jejichž krajní body jsou čísla, kterým se přisuzuje z různých důvodů zvláštní význam.

Opět srovnáme obě metody; jako počet vyčíslení zvolíme (pro všechny případy) např. $N = 20$, ať se pokud možno vyhneme „patologickým“ jevům. U MPI bude znovu $\delta = 0.02$.

$I_0 = [-e, \pi]$	MZŘ	MPI
\bar{x}	-1.02854639945572	-1.03018303892649
$ \bar{x} - x^* $	0.000107327664302437	0.00174396713507585
$f(\bar{x})$	7.58071261534158	7.58073199638867
poslední ILM	$[-1.02885979524217, -1.02823300366927]$	$[-1.05302477451343, -1.00734130333955]$

Výchozí ILM, mající jako krajní body nejznámější matematické konstanty, znovu přinesl vyšší přesnost v případě MZŘ. Přesto bychom u MZŘ asi čekali přesnější výsledek, protože pro mnohem širší $I_0 = [-10, 9]$ dosáhla v předchozím MZŘ pro $N = 20$ přesnosti (resp. chyby) řádově asi $4 \cdot 10^{-5}$.

$I_0 = [-\varphi, \varphi]$	MZŘ	MPI
\bar{x}	-1.02840368399006	-1.02842226424431
$ \bar{x} - x^* $	$3.53854416286215e - 05$	$1.68051873727926e - 05$
$f(\bar{x})$	7.58071254973216	7.58071254353649
poslední ILM	$[-1.02857675426177, -1.02823061371835]$	$[-1.04998284431145, -1.00686168417718]$

„Hezký“ výchozí ILM symetrický kolem nuly, jenž má délku dvojnásobku zlatého řezu, „paradoxně“ vyšel v neprospěch metody zlatého řezu, což se nám dosud nestalo (ačkoli rozdíl v přesnosti obou metod je minimální).

$I_0 = [-666, 13]$	MZŘ	MPI
\bar{x}	-1.03675476692745	-1.2755078125
$ \bar{x} - x^* $	0.00831569717320879	0.247068742745762
$f(\bar{x})$	7.58115673754269	8.03635937956143
poslední ILM	$[-1.07306881079478, -1.00044072306011]$	$[-1.62703125, -0.923984374999999]$

„Zakletý“ $I_0 = [-666, 13]$ kupodivu nedal tak nešťastné výsledky, jak by se vzhledem k jeho šířce a „vychýlenosti“ dalo čekat. Zejména MZŘ se díky vyššímu počtu dostupných ILM celkem solidně přiblížila bodu skutečného minima $x^* = 1.028490691$.

$I_0 = [-7, 69]$	MZŘ	MPI
\bar{x}	-1.02638557966148	-1.00853515625
$ \bar{x} - x^* $	0.00205348885284229	0.0199039122643225
$f(\bar{x})$	7.58073944912128	7.58321160260756
poslední ILM	$[-1.03045018545517, -1.02232097386779]$	$[-1.065625, -0.9514453125]$

A nakonec „šťastný“ výchozí ILM, u něhož – jak bývá zvykem – dala přesnější výsledky MZŘ.

I když je $I_0 = [-7, 69]$ o dost užší než předchozí $I_0 = [-666, 13]$, dosažená přesnost se zas tak podstatně neliší. Chování metod je pro takto velká N z velké míry nepředvídatelné.

Analýza numerických metód pre funkciu viacerých premenných

Základné informácie

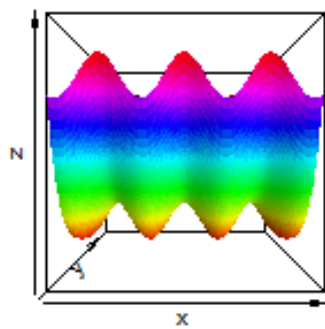
V tejto časti sa zameriame na analýzu numerických metód pre funkciu

$$f(x, y) = \cos(x) + y^2 + y^4.$$

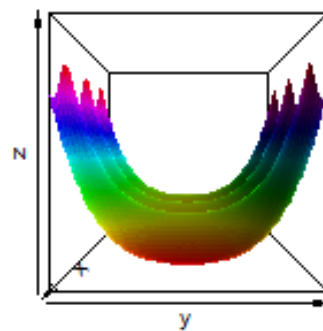
Na ďalšej strane je táto funkcia z rôznych pohľadov spolu s grafom znázorňujúcim vrstevnice funkcie. „Presné“ minimum funkcia nadobúda v bode $[x^*, y^*] = [(2k - 1)\pi, 0]$ pre $k \in \mathbb{Z}$. Hodnota tohto minima je $f(x^*, y^*) = -1 = f^* = \inf_{[x, y] \in \mathbb{R}^2} f(x)$.

Metódy, ktorými budeme skúmať túto funkciu sú **metóda najväčšieho spádu** a **metóda združených gradientov**.

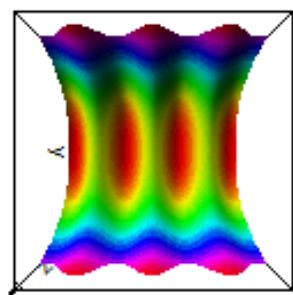
Pohľad na rovinu xz



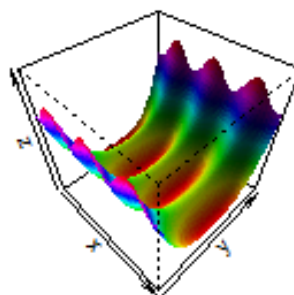
Pohľad na rovinu yz



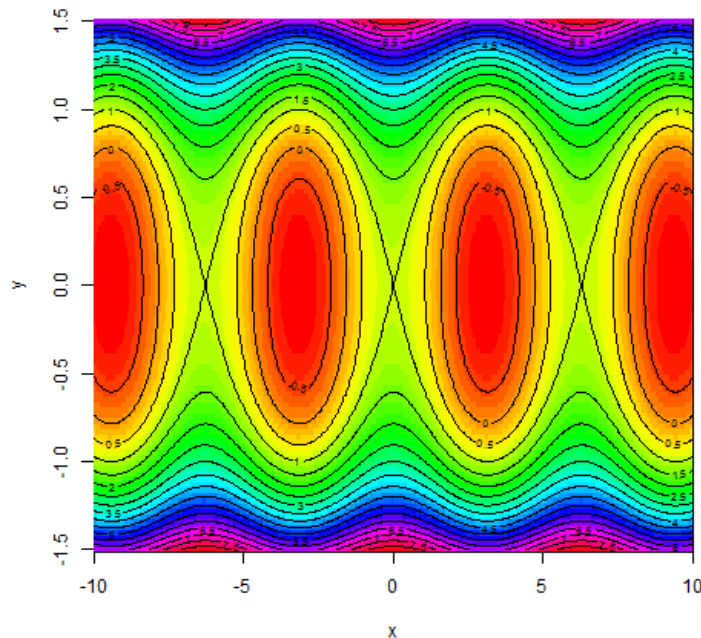
Pohľad na rovinu xy



Pohľad v \mathbb{R}^3



Obrázek 1.2: Graf funkcie $f(x, y) = \cos(x) + y^2 + y^4$



Obrázek 1.3: Vrstevnice funkcie $f(x, y) = \cos(x) + y^2 + y^4$

1.3 Metóda najväčšieho spádu

1.3.1 Teória

V tejto časti sa zameriame na analýzu metódy najväčšieho spádu. Ide o metódu 1. rádu, ktorá je založená na vytvorení minimalizujúcej postupnosti $\{x^{[k]}\}$ (ide o spádovú metódu), kde jednotlivé členy počítame ako

$$x^{[k+1]} = x^{[k]} + \alpha_k h_k.$$

Smer k -teho kroku určíme ako

$$h_k = -\text{grad } f(x^{[k]}).$$

Ide teda o gradientnú metódu.

Dĺžku k -teho kroku určíme na základe presnej minimalizácie

$$f(x^{[k]} - \alpha_k \text{grad } f(x^{[k]})) = \min_{\alpha \geq 0} f(x^{[k]} - \alpha \text{grad } f(x^{[k]})).$$

Tento postup si ukážeme pre našu funkciu $f(x, y) = \cos(x) + y^2 + y^4$ a počiatočnú aproximáciu $x^{[0]} = [\pi, 1]$:

$$\text{grad } f(x, y) = \begin{pmatrix} -\sin(x) \\ 2y + 4y^3 \end{pmatrix},$$

$$h_0 = -\text{grad } f(x^{[0]}) = -\begin{pmatrix} -\sin(\pi) \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \end{pmatrix},$$

$$f(x^{[0]} + \alpha h_0) = f(\pi, 1 - 6\alpha)$$

$$-1 + (1 - 6\alpha)^2 + (1 - 6\alpha)^4 \rightarrow \min_{\alpha \geq 0}$$

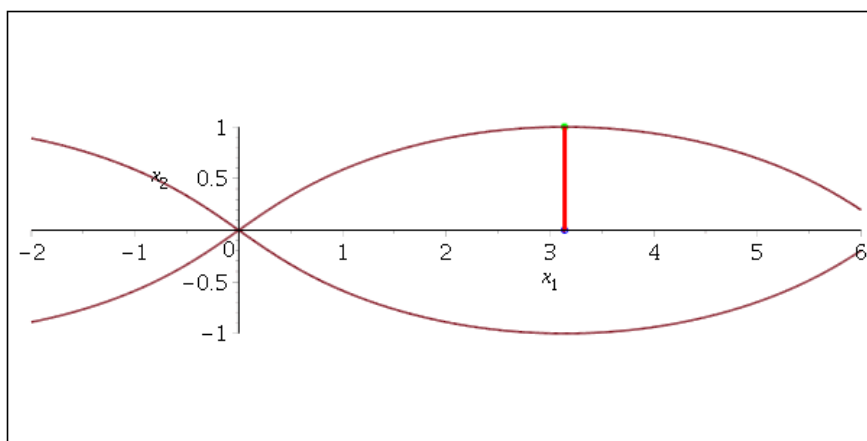
deriváciou podľa α získame

$$(1 - 6\alpha) \cdot (-12 - 24(1 - 6\alpha)^2) = 0.$$

Riešením je $\alpha = \frac{1}{6}$.

$x^{[1]} = \begin{pmatrix} \pi \\ 1 \end{pmatrix} + \frac{1}{6} \cdot \begin{pmatrix} 0 \\ -6 \end{pmatrix} = \begin{pmatrix} \pi \\ 0 \end{pmatrix}$, pričom tento bod je priamo bod minima $f(\pi, 0) = -1$.

Na nasledujúcom obrázku vidíme grafické znázornenie.



Obrázek 1.4: Grafické znázornenie MNS pre počiatkový bod $[\pi, 1]$

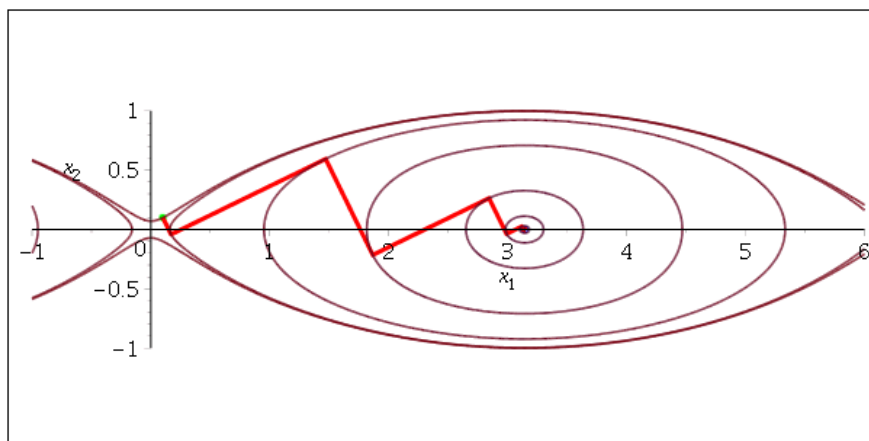
1.3.2 Analýza metódy pri rôznych hodnotách počiatočnej aproximácie

V tejto časti sa pozrieme na to, ako sa metóda najväčšieho spádu správa pre rôzne počiatočné body. Ako ukončovacie kritérium pre porovnanie jednotlivých bodov si zvolíme $|f(x^{[k+1]}) - f(x^{[k]})| < \epsilon$, kde $\epsilon = 0.001$.

Bod $[0.1, 0.1]$

počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	$[0.1690128036, -0.0410210369]$	0.9874368599	1.98743686
5	$[2.989936858, -0.0368663233]$	-0.9871613106	0.0128386894
7 (ϵ)	$[3.130529833, -0.00270686397]$	-0.9999314804	6.85196×10^{-5}
10	$[3.141463117, 0.0001323478754]$	-0.9999999741	2.59×10^{-8}
13	$[3.141588471, -0.00000102342862]$	-1	0

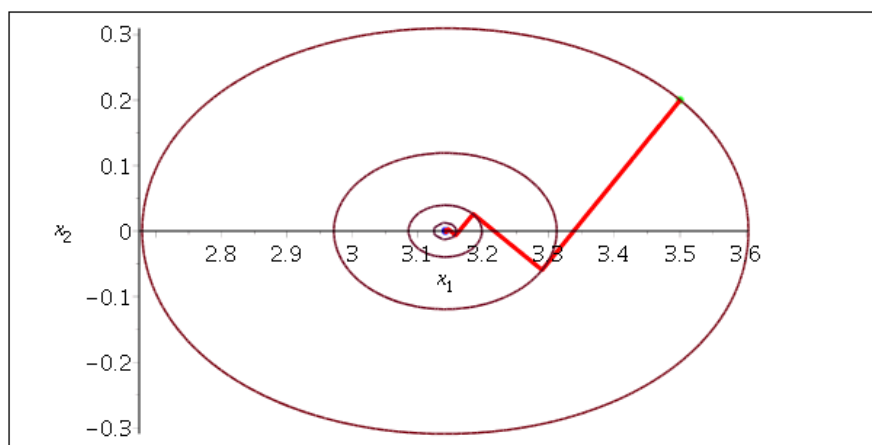
Táto tabuľka zachytáva niekoľko bodov minimalizujúcej postupnosti, funkčné hodnoty v týchto bodoch a rozdiel tejto funkčnej hodnoty oproti hodnote našej úlohy. Konkrétne v 13. iterácii dospievame do bodu, v ktorom je funkčná hodnota rovná -1, čiže hodnote úlohy (aj na základe zaokrúhľovacích chýb). Na splnenie nášho ukončovacieho kritéria je potrebných 7. iterácií.



Obrázek 1.5: Grafické znázornenie MNS pre počiatočný bod $[0.1, 0.1]$

Bod [3.5, 0.2]

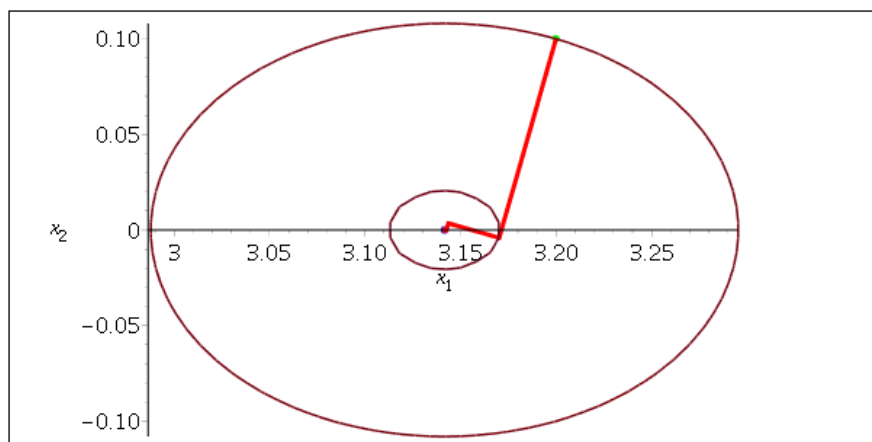
počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	[3.289392237, -0.0593697378]	-0.9855603203	0.0144396797
4 (ϵ)	[3.146145970, 0.002803718632]	-0.9999817728	1.82272×10^{-5}
5	[3.143304884, -0.000695163115]	-0.9999980508	1.9492×10^{-6}
10	[3.141598221, 0.00000342799357]	-1	0



Obrázek 1.6: Grafické znázornenie MNS pre počiatkový bod [3.5, 0.2]

Bod [3.2, 0.1]

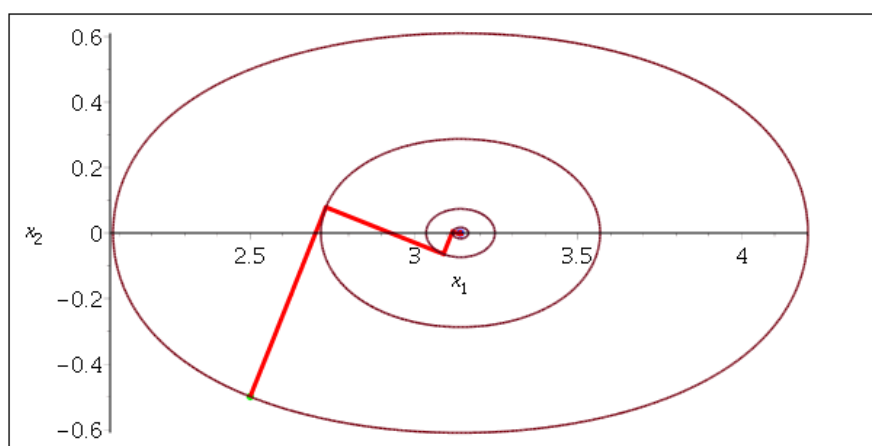
počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	[3.170213671, -0.0040942237]	-0.9995736836	4.263164×10^{-4}
2 (ϵ)	[3.143606773, 0.003519280466]	-0.9999855862	1.44138×10^{-5}
5	[3.141625350, -0.0000046780440]	-0.9999999995	5×10^{-10}
6	[3.141594954, 0.000004019753136]	-1	0



Obrázek 1.7: Grafické znázornenie MNS pre počiatkový bod $[3.2, 0.1]$

Bod $[2.5, -0.5]$

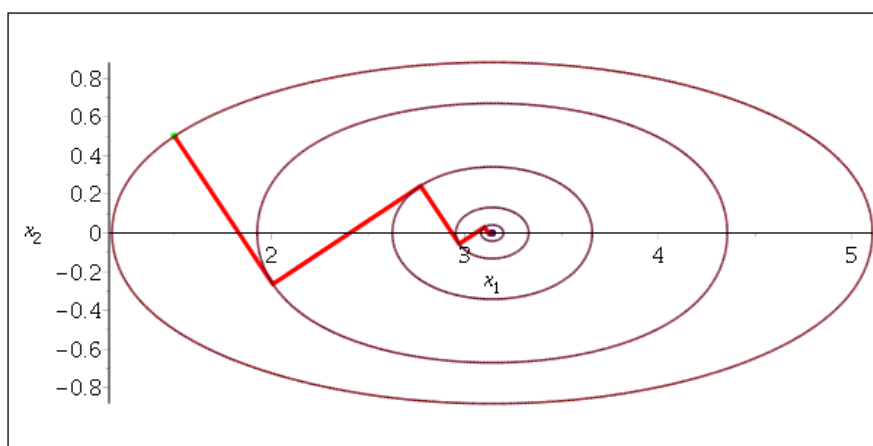
počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	$[2.730880636, 0.0786751447]$	-0.9106086817	0.0893913183
4 (ϵ)	$[3.138664285, -0.003669701672]$	-0.9999822454	1.77546×10^{-5}
5	$[3.140236378, 0.000270564286]$	-0.9999990071	9.929×10^{-7}
9	$[3.141588412, 8.4606731 \times 10^{-7}]$	-1	0



Obrázek 1.8: Grafické znázornenie MNS pre počiatkový bod $[2.5, -0.5]$

Bod $[1.5, 0.5]$

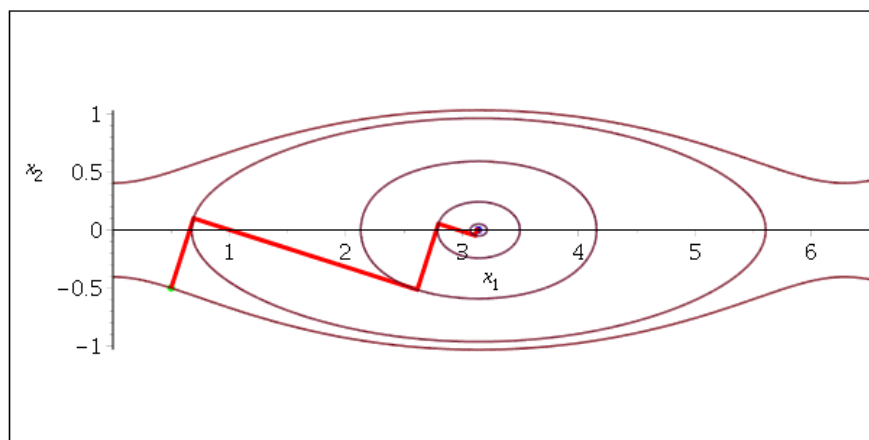
počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	$[2.008240241, -0.2642748801]$	-0.3489064461	0.6510935539
5	$[3.124961325, -0.00554259140]$	-0.9998303154	1.696846×10^{-4}
6 (ϵ)	$[3.137679941, 0.002941845463]$	-0.9999836907	1.63093×10^{-5}
10	$[3.141556527, 0.00002716313187]$	-0.9999999986	1.4×10^{-9}
12	$[3.141589182, -0.000002610062399]$	-1	0



Obrázek 1.9: Grafické znázornenie MNS pre počiatkový bod $[1.5, 0.5]$

Bod $[0.5, -0.5]$

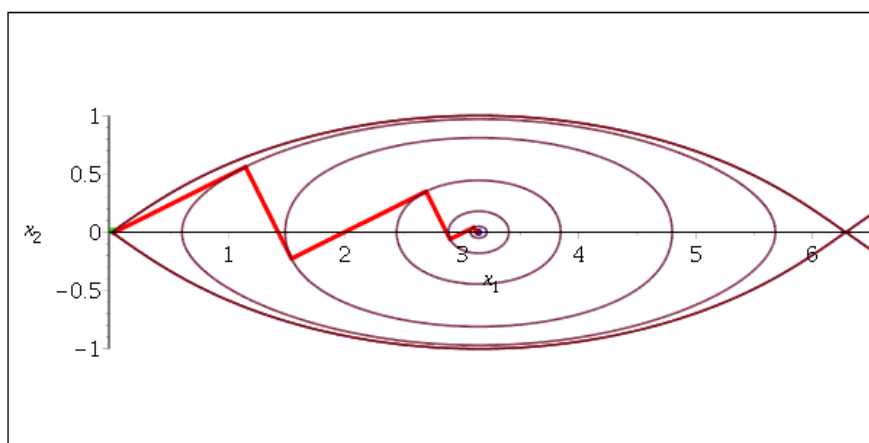
počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	$[0.6917520783, 0.0999432535]$	0.7802179959	1.780217996
5	$[3.127175990, 0.00230380136]$	-0.9998907742	1.092258×10^{-4}
6 (ϵ)	$[3.140369704, -0.001913134278]$	-0.9999955921	4.4079×10^{-6}
10	$[3.141590662, -0.000003115015096]$	-1	0



Obrázek 1.10: Grafické znázornenie MNS pre počiatočný bod $[0.5, -0.5]$

Bod $[0.01, 0.01]$

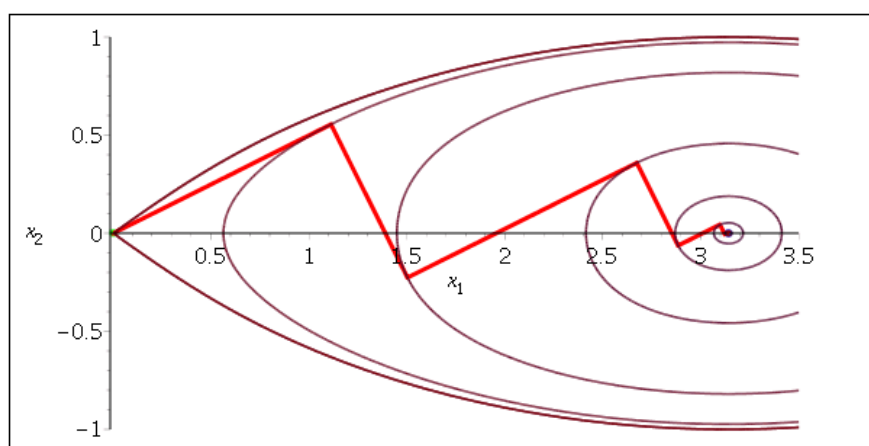
počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	$[0.01714034889, -0.00428379198]$	0.9998714590	1.999871459
5	$[2.896530382, -0.0602046099]$	-0.9664844859	0.0335155141
8 (ϵ)	$[3.138497576, 0.003095684120]$	-0.9999856269	1.43731×10^{-5}
10	$[3.141363439, 0.0002292643358]$	-0.9999999211	7.89×10^{-8}
13	$[3.141585109, -0.00000188573662]$	-1	0



Obrázek 1.11: Grafické znázornenie MNS pre počiatočný bod $[0.01, 0.01]$

Bod $[0.001, 0.001]$

počet iterácií (n)	bod minima $x^{[n]}$	funkčná hodnota $f(x^{[n]})$	$ f(x^{[n]}) - f^* $
1	$[0.001714283205, -0.000428569505]$	0.9999987143	1.9999987143
5	$[2.884356602, -0.0630995246]$	-0.9630994406	0.0369005594
8 (ϵ)	$[3.138334523, 0.003258062876]$	-0.9999840772	1.59228×10^{-5}
10	$[3.141351305, 0.0002413487012]$	-0.9999999127	8.73×10^{-8}
13	$[3.141584738, -0.00000198639614]$	-1	0



Obrázek 1.12: Grafické znázornenie MNS pre počiatkový bod $[0.001, 0.001]$

Bod $[0, 1]$

V tomto bode metóda zlyháva, preto si ukážeme postup výpočtu.

$$\text{Smer 1. kroku } h_0 = - \begin{pmatrix} -\sin(0) \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \end{pmatrix}$$

$$f(x^{[0]} + \alpha h_0) = f(0, 1 - 6\alpha)$$

$$1 + (1 - 6\alpha)^2 + (1 - 6\alpha)^4 \rightarrow \min_{\alpha \geq 0}$$

Po zderivovaní a položení $= 0$ zistíme, že dĺžka kroku je $\alpha_0 = \frac{1}{6}$. Ďalší bod minimalizujúcej postupnosti je teda $x^{[1]} = x^{[0]} + \alpha_0 h_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{6} \cdot \begin{pmatrix} 0 \\ -6 \end{pmatrix} = [0, 0]$ a hodnota funkcie v tomto bode je $f(0, 0) = 1$. Počítajme ďalej smer nasledujúceho kroku:

$h_1 = - \begin{pmatrix} -\sin(0) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Ide teda o stacionárny bod, ale jedná sa o lokálne maximum tejto funkcie. Metóda tento bod vyhodnotí ako minimum a ďalej sa už z 0 neodrazí.

Bod	Počet iterácií (ϵ)	Počet iterácií ($f(x^{[n]}) = -1$)
$[\pi, 1]$	1	1
$[0.1, 0.1]$	7	13
$[3.5, 0.2]$	4	10
$[3.2, 0.1]$	2	6
$[2.5, -0.5]$	4	9
$[1.5, 0.5]$	6	12
$[0.5, -0.5]$	6	10
$[0.01, 0.01]$	8	13
$[0.001, 0.001]$	8	13
$[0, 1]$	zlyháva	zlyháva

Tabuľka obsahuje zhrnutie skúmaných bodov. V 2. stĺpci sa nachádza počet iterácií potrebných na splnenie ukončovacieho kritéria a v 3. stĺpci je počet iterácií potrebných na to, aby bola hodnota v danom bode zaokrúhlená na skutočnú hodnotu minima $\rightarrow -1$.

Vidíme, že len v prípade, keď sa nachádzame priamo nad, resp. pod bodom minima, tj. $[x^*, y^*] = [(2k-1)\pi, 0]$ pre $k \in \mathbb{Z}$, metóda dospeje do tohto minima hneď v 1. iterácii. Čím viac sa vzdďaľuje od bodu minima na obe strany, tým je metóda pomalšia, ale správa sa štandardne. Výnimkou sú body, ktorých x -ová súradnica má tvar $2k\pi$ pre $k \in \mathbb{Z}$. V týchto bodoch metóda zlyháva, keďže v 1. iterácii dospeje do bodu $[2k\pi, 0]$, vyhodnotí tento bod ako minimum a ďalej nepokračuje. Tieto body sú ale bodmi lokálneho maxima.

Pre body v blízkosti bodov lokálneho maxima platí, že metóda skončí na základe ukončovacieho kritéria po 6-8 iteráciách s tým, že rozdiel hodnoty funkcie v spočítaných bodoch a skutočnej hodnoty minima ($|f(x^{[n]}) - f^*|$) je rádovo približne 1×10^{-5} . Metóda najväčšieho spádu je teda veľmi pomalou metódou.

1.4 Metóda združených gradientov

1.4.1 Teória

$$f(x, y) = \cos(x) + y^2 + y^4 \rightarrow \min$$

Aplikovanie metódy združených gradientov pre nekvadratickú funkciu je rovnaké ako pre kvadratickú funkciu, no s odlišnou voľbou β_k . Postupnosť $g_k := \text{grad } f(x^{[k]})$ nemusí konvergovať v n -krokoch, ale pokiaľ sú splnené predpoklady:

$f \in C^1$ na \mathbb{R}^n a $x^{[0]} \in \mathbb{R}^n$ také, že je množina $\{x \in \mathbb{R}^n \mid f(x) \leq f(x^{[0]})\}$ ohraničená

MSG sa môže po n -krokoch reštartovať. Čo znamená, že zoberieme za nové $x^{[1]}$ výsledok MSG s cyklom dĺžky n a následne $x^{[2]}$ je výsledkom ďalšieho cyklu s počiatočným bodom

$x^{[1]}$ atď. Dôsledkom reštartu je rýchlejšia konvergencia.

Teoretický postup pri výpočte jednotlivých krokov: MSG určuje bod $x^{[k+1]}$ pomocou minimalizácie funkcie f v smere vektoru h_k

$$x^{[k+1]} = x^{[k]} + \alpha_k h_k = x^{[0]} + \sum_{i=0}^k \alpha_i h_i$$

Platia vzťahy

$$h_0 = -g_0 = -\text{grad } f(x^{[0]}); \quad h_k = -g_k + \beta_{k-1} h_{k-1} = -\text{grad } f(x^{[k]}) + \beta_{k-1} h_{k-1}$$

$$x^{[k+1]} = x^{[k]} + \alpha_k h_k = x^{[0]} + \sum_{i=0}^k \alpha_i h_i$$

$$\alpha_k = -\frac{h_k^T \text{grad } f(x^{[k]})}{h_k^T H_k h_k}$$

kde H_k je Hessian v bode $x^{[k]}$.

Použitie značenie $x^{[k]} = [\tilde{x}; \tilde{y}]$

V našom konkrétnom príklade platí:

$$f'_x = -\sin(x); \quad f'_y = 2y + 4y^3$$

$$\text{grad } x^{[k]} = \begin{pmatrix} -\sin(\tilde{x}) \\ 2\tilde{y} + 4\tilde{y}^3 \end{pmatrix}$$

$$H(x^{[k]}) = \begin{pmatrix} -\cos(\tilde{x}) & 0 \\ 0 & 2 + 12\tilde{y}^2 \end{pmatrix}$$

Voľba β_k má narozdiel od kvadratické funkcie viacero možností:

$$\begin{aligned} \beta_k^{FR} &:= \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} & \beta_k^{PR} &:= \frac{(g_k - g_{k-1})^T g_k}{g_{k-1}^T g_{k-1}} \\ \beta_k^{HS\#1} &:= \frac{(g_k - g_{k-1})^T g_k}{h_{k-1}^T (g_k - g_{k-1})} & \beta_k^{HS\#2} &:= \frac{(g_k - g_{k-1})^T g_k}{g_{k-1}^T (g_k - g_{k-1})} \end{aligned}$$

1.4.2 Analýza chovania metódy pri rôznych hodnotách počiatočného bodu

Zvolená funkcia

$$f(x, y) = \cos(x) + y^2 + y^4$$

nemá v \mathbb{R}^n jedno globálne minimum. Množina bodov $\{[x^*, y^*] = [(2k - 1)\pi, 0]\}$ pre $k \in \mathbb{Z}$, ktoré nadobúdajú hodnotu minima $f(x^*, y^*) = -1$, je neobmedzená.

Zameriame sa na analýzu chovania metódy pri rôznych počiatočných hodnotách na obmedzenej množine M lokalizácie minima

$$M = \{[x, y] \mid x \in [0; 6]; y \in [-5; 5]\}$$

Táto množina obsahuje práve jedno minimum a to v bode $[\pi, 0]$

Bod $[0, y]$

Voľbu počiatočnej hodnoty x -ovej súradnice 0 spôsobuje zlyhanie metódy. Metóda po prvej iterácii prehlási nájdenie minima v bode $[0; 0]$ s hodnotou funkcie $\bar{x} = 1$ a ďalej v iteráciách nepokračuje. Hodnota gradientu v bode $[0; 0]$ je $(0, 0)$, preto je tento bod považovaný za minimum a metóda skončí.

Vybrané body s rýchlou konvergenciou

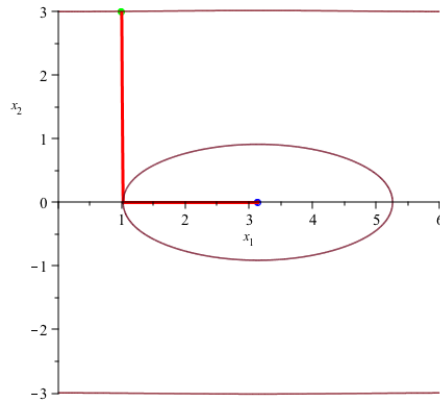
Výber niektorých počiatočných bodov $x^{[0]}$ z množiny M , pre ktoré je nutný počet iterácií pre nájdenie minima nízky, pretože MSG konverguje veľmi rýchlo. Pri použití resetu β sa rýchlosť konvergenzie nezmení alebo sa ešte zvýši. Dôvodom je počiatočný nízky počet nutných iterácií.

$\mathbf{x}^{[0]}$	počet iterácií bez resetu	počet iterácií s resetom
$x^{[0]} = [6; 5]$	3	3
$x^{[0]} = [6; -5]$	3	3
$x^{[0]} = [3; 1]$	4	3
$x^{[0]} = [1; 3]$	4	3
$x^{[0]} = [0.5; 0]$	1	1
$x^{[0]} = [0.5; 5]$	3	3
$x^{[0]} = [0.05; 5]$	2	2
$x^{[0]} = [0.005; 5]$	2	2

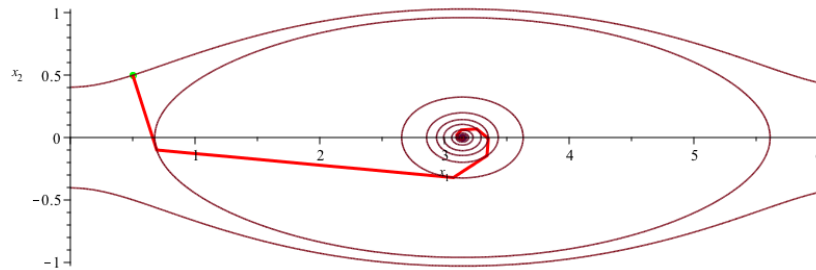
Na obrázku 1.13 je graficky znázornená MSG pre počiatočný bod $[1, 3]$ bez resetu β .

Bod $[0.5, 0.5]$

Pri počiatočnom bode $x^{[0]} = [0.5, 0.5]$ je pekne vidieť zmenu rýchlosti konvergenzie pri použití metódy s a bez resetu β



Obrázek 1.13: Grafické znázornenie MSG pre počiatkový bod $[1,3]$ bez resetu β



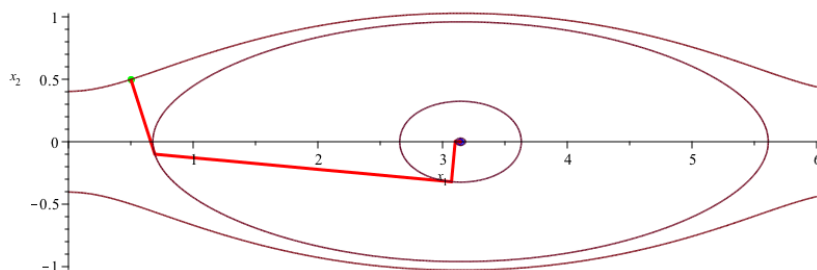
Obrázek 1.14: Grafické znázornenie MSG pre počiatkový bod $[0.5,0.5]$ bez resetu β

$\mathbf{x}^{[0]}$	počet iterácií	\bar{x}	$ \bar{x} - x^* $
$x^{[0]} = [0.5; 0.5]$	5	-0.9886965275	$1.13034725x10^{-2}$
$x^{[0]} = [0.5; 0.5]$	10	-0.9998901505	$1.098495x10^{-4}$
$x^{[0]} = [0.5; 0.5]$	20	-0.9999999721	$2.79x10^{-8}$
$x^{[0]} = [0.5; 0.5]$	27	-1	0
$x^{[0]} = [0.5; 0.5]$	6 s resetom β	-1	0

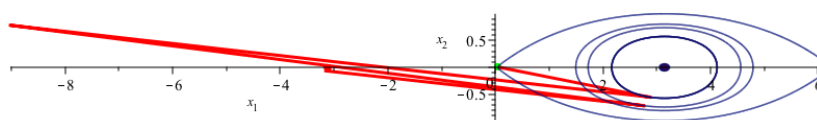
Z tabuľky je zjavné, že pri použití resetu β sa konvergencia metódy viacnásobne zvýši. Graficky znázornené iteračné kroky je vidno na obrázku 1.14 a 1.15.

Bod $[0.05, 0.005]$

$\mathbf{x}^{[0]}$	počet iterácií	\bar{x}	$ \bar{x} - x^* $
$x^{[0]} = [0.05; 0.005]$	5	$-0,1516818598$	$0,8483181402$
$x^{[0]} = [0.05; 0.005]$	10	-0.9979485357 ($x^{[10]}$ leží mimo def. obor)	$2.0514643x10^{-3}$
$x^{[0]} = [0.05; 0.005]$	20	-0.9996927007 ($x^{[20]}$ leží mimo def. obor)	$3.072993x10^{-4}$
$x^{[0]} = [0.05; 0.005]$	50	-0.9999991770 ($x^{[50]}$ leží mimo def. obor)	$8.23x10^{-7}$
$x^{[0]} = [0.05; 0.005]$	6 s resetom β	-1	0



Obrázek 1.15: Grafické znázornenie MSG pre počiatkový bod $[0.5,0.5]$ s resetom β

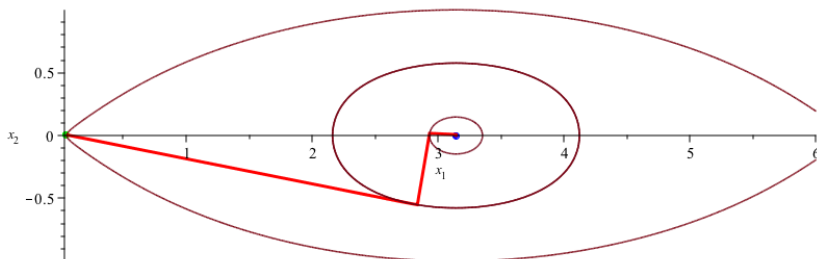


Obrázek 1.16: Grafické znázornenie MSG pre počiatkový bod $[0.05;0.005]$ bez resetu β

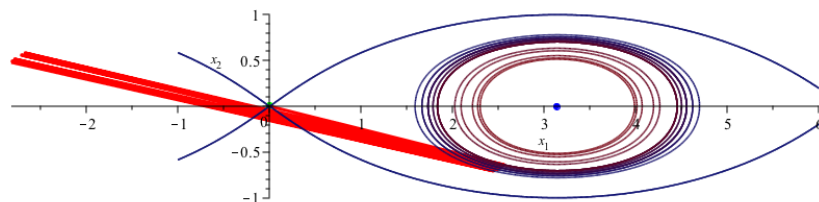
Pre počiatkový bod $[0.05;0.005]$ a počet iterácií 50 sú graficky znázornené kroky na obrázku 1.16. Chyba aproximácie je síce nízka $8,23 \times 10^{-7}$ avšak lokalizácia minima je v bode $[-3.141590818; -0.00090719658453]$, čo je síce minimum predpísanej funkcie f pre voľbu $k = 0$, ale je mimo nami definovanú množinu hodnôt M . MSG opustí množinu lokalizácie minima v 4. iterácii, potom sa do množiny vráti v 5. a opäť množinu opustí. Takto sa vracia tam a späť a limitne smeruje k minimu v bode $[-\pi, 0]$. Pokiaľ resetujeme parameter β , tak metóda nájde minimum už v 6. iterácii a jedná sa o minimum v množine M . Grafické znázornenie na obrázku 1.17.

Bod $[0.005, 0.005]$

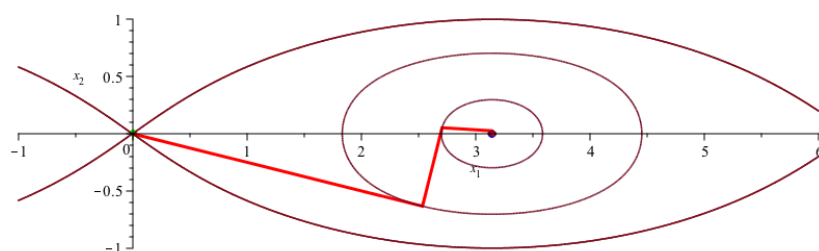
Pre počiatkový bod $[0.005;0.005]$ MSG zlyháva, pretože nekonverguje v cykle dĺžky n . Graficky znázornené na obrázku 1.18. S resetom β je metóda účinná a nájde minimum už v 6. iterácii. Grafické znázornenie na obrázku 1.19.



Obrázek 1.17: Grafické znázornenie MSG pre počiatkový bod $[0.05;0.005]$ s resetom β



Obrázek 1.18: Grafické znázornenie MSG pre počiatočný bod $[0.005; 0.005]$ bez resetu β



Obrázek 1.19: Grafické znázornenie MSG pre počiatočný bod $[0.005; 0.005]$ s resetom β

Všeobecne neviem určiť chybné správanie metódy na určitom intervale. Nejednoznačnosť konvergenencie alebo zmena množiny konvergenencie nastáva v δ -okolí ($|\delta| > 0$) množiny bodov $\{[x, y] = [2k\pi; 0]; k \in \mathbb{Z}\}$

Existujú 3 scenáre pre vybrané počiatočné body:

1. MSG so zvoleným počiatočným bodom konverguje na vybranom intervale lokalizácie minima
2. MSG so zvoleným počiatočným bodom konverguje k bodu, ktorý ale nie je v intervale lokalizácie minima. MSG s resetom β urýchli konvergenciu, avšak minimum stále nemusí patriť do intervale lokalizácie minima.
3. MSG so zvoleným počiatočným bodom nekonverguje v cykle dĺžky n . Ale MSG s resetom β upraví konvergenciu a funkcia s rovnakým počiatočným bodom konverguje.

V nižšie uvedenej tabuľke je ilustrovaná konvergenция metódy k bodu minima bez resetu β v porovnaní s konvergeniou s resetom β pre pár vybraných počiatočných bodov. Metóda sa správa podľa vyššie uvedeného scenára č.2, t.j. po resete nastáva zmena lokalizácie minima.

$\mathbf{x}^{[0]}$	bod minima bez resetu β	bod minima po resete β
$x^{[0]} = [0.1, 0.1]$	$[-\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.1, 0.2]$	$[3\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.2, 0.2]$	$[-3\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.3, 0.1]$	$[-9\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.3, 0.3]$	$[3\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.4, 0.1]$	$[-7\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.4, 0.2]$	$[-9\pi, 0]$	$[-\pi, 0]$
$x^{[0]} = [0.4, 0.3]$	$[3\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.5, 0.1]$	$[5\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.5, 0.2]$	$[-11\pi, 0]$	$[-\pi, 0]$
$x^{[0]} = [0.6, 0.1]$	$[7\pi, 0]$	$[3\pi, 0]$
$x^{[0]} = [0.6, 0.2]$	$[13\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.6, 0.3]$	$[3\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.7, 0.1]$	$[9\pi, 0]$	$[5\pi, 0]$
$x^{[0]} = [0.7, 0.2]$	$[-\pi, 0]$	$[-\pi, 0]$
$x^{[0]} = [0.7, 0.3]$	$[3\pi, 0]$	$[\pi, 0]$
$x^{[0]} = [0.8, 0.2]$	$[-\pi, 0]$	$[-\pi, 0]$
$x^{[0]} = [0.8, 0.3]$	$[-\pi, 0]$	$[-\pi, 0]$
$x^{[0]} = [0.9, 0.2]$	$[-5\pi, 0]$	$[-5\pi, 0]$
$x^{[0]} = [0.9, 0.3]$	$[-\pi, 0]$	$[-\pi, 0]$
$x^{[0]} = [1, 0.3]$	$[-\pi, 0]$	$[-\pi, 0]$