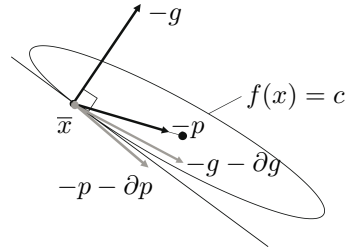


Robustnost gradientního směru

Důležitou vlastností gradientních metod je jejich robustnost. Z obr. 4.5 je patrné, že i když ∂g je relativně velká perturbace gradientu g , vektor $-g - \partial g$ je stále směrem poklesu, zatímco malá perturbace ∂p optimálního směru p může způsobit, že $-p - \partial p$ není ani směrem poklesu.



Obr. 4.5: Robustnost gradientu a směru maximálního poklesu g a p

4.3 Aproximace gradientu pomocí konečných diferencí

Některé důležité cenové funkce jsou diferencovatelné, avšak není možno zadat explicitní pravidlo pro výpočet gradientu. Jako příklad si uveďme typickou úlohu tvarové optimalizace v mechanice, kde $x \in \mathbb{R}^n$ jsou tzv. *návrhové proměnné* popisující tvar tělesa, ale cenový funkcionál f je zadán pomocí funkce

$$J : \mathbb{R}^{n_s} \rightarrow \mathbb{R}$$

definované pro tzv. *stavové proměnné* $u \in \mathbb{R}^{n_s}$, které splňují stavové rovnice

$$K(x)u = f(x).$$

V mechanice mohou být stavové proměnné posunutí a výše uvedená rovnice je rovnice rovnováhy definovaná pozitivně definitní maticí $K(x)$ a vektorem $f(x) \in \mathbb{R}^{n_s}$. K výpočtu hodnoty cenové funkce

$$f(x) = J(u(x))$$

pro dané x je tak třeba najít řešení $u = u(x)$ poměrně rozsáhlé (tisíce až miliony neznámých) soustavy rovnic rovnováhy. Návrhových proměnných bývá obvykle mnohem méně, řádově desítky až stovky.

Konečné difference

První derivace potřebné k vyčíslení gradientu mohou být nejsnáze vyčísleny pomocí *dopředných diferencí*

$$\frac{\partial f(x^k)}{\partial x_i} \approx \frac{1}{\varepsilon} (f(x^k + \varepsilon e_i) - f(x^k)), \quad (4.7)$$

kde e_i je i -tý sloupec jednotkové matice a $\varepsilon > 0$ je malé číslo. V přesné aritmetice by platilo čím menší, tím lepší, avšak při praktické počítačové implementaci je třeba toto pravidlo sladit s faktem, že čím menší ε , tím větší je vliv zaokrouhlovacích chyb způsobených počítačovou aritmetikou.

Dopředné difference jsou přesné pouze pro lineární funkce. Vyšší přesnosti je možno dosáhnout s *centrálními diferencemi*

$$\frac{\partial f(x^k)}{\partial x_i} \approx \frac{1}{2\varepsilon} (f(x^k + \varepsilon e_i) - f(x^k - \varepsilon e_i)), \quad (4.8)$$

kteřé jsou přesné pro kvadratické polynomy, avšak cena výpočtu je dvojnásobná. V řadě případů lze využít speciálního tvaru funkce k efektivnějšímu implicitnímu výpočtu hodnoty gradientu.

4.4 Ukončovací podmínky

Jako ukončovací podmínku používáme v této kapitole velikost gradientu, tj. $\|g^k\| = \|\nabla f(x^k)\|$. Tímto testem ověřujeme, zda je bod x^k dostatečně přesnou aproximací stacionárního bodu funkce f . Alternativní ukončovací podmínkou může být test na velikost kroku, tj. $\|x^{k+1} - x^k\|$. Jinou možností je ukončovat iterační minimalizační proces při dostatečně malé změně funkční hodnoty, tj. $\|f(x^{k+1}) - f(x^k)\|$.

Vhodnou ukončovací podmínkou je také test na relativní velikost gradientu.

Příklady k procvičení

1. Nakreslete graf a vrstevnice funkce

$$f(x_1, x_2) = x_1^2 + x_2^2.$$

Řešte úlohu optimalizace bez omezení

$$\min_{x \in \mathbb{R}^2} f(x_1, x_2).$$

Tuto úlohu řešte metodou největšího spádu s analyticky vypočteným gradientem a poté metodou největšího spádu s numericky vypočteným gradientem. Úlohu počítejte s ukončující podmínkou na velikost gradientu i délku kroku.