Metody podmíněné optimalizace

- Stále se pohybujeme v doméně [striktně] konvexní optimalizace
- Řešíme tzv. *úlohu matematického programování ve standardním tvaru*:

$$\begin{aligned} &\min f(\mathbf{x}) \\ g_i(\mathbf{x}) &\leq 0, \qquad i = 1, 2, \ldots, n \\ h_j(\mathbf{x}) &= 0, \qquad j = 1, 2, \ldots, m \end{aligned}$$

- Konvexita $\Leftrightarrow$ f a g_i jsou konvexní, h_j jsou lineární
- Dostupné "polynomiální" algoritmy pracují s touto formou zadání problému
- Cílem je převést úlohu na jednodušší, v ideálním případě bez omezujících podmínek

Pouze omezení ve tvaru rovností

- Žádné nerovnosti $m = 0$
- Přeformulování úlohy

$$
\begin{aligned}
\min f(\mathbf{x}) & \\
h_j(\mathbf{x}) = 0, & \qquad j = 1, 2, \ldots, m
\end{aligned}
$$

$$\Updownarrow$$

$$\min L(\mathbf{x}, \lambda) \quad = \quad f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j h_j(\mathbf{x})$$

- L je tzv. *Lagrangeova* funkce a λ jsou Lagrangeovy multiplikátory
- Převedeno na problém nepodmíněné optimalizace, ale za cenu přidání m proměnných

Karush-Kuhn-Tuckerovy podmínky

- Funkce f, g_i a h_j opět v okolí optima $\mathbf{x}_{opt}$ rozvineme do Taylorovy řady
- Cílová funkce

$$\begin{aligned} f(\mathbf{x}) &\approx f(\mathbf{x}_{opt}) + \nabla\mathbf{f}(\mathbf{x}_{opt})^{\mathsf{T}}(\mathbf{x}-\mathbf{x}_{opt}) \geq f(\mathbf{x}_{opt}) \Rightarrow \\ \nabla\mathbf{f}(\mathbf{x}_{opt})^{\mathsf{T}}(\mathbf{x}-\mathbf{x}_{opt}) &\geq 0 \end{aligned}$$

- Omezení ve tvaru rovností

$$\begin{aligned} h_j(\mathbf{x}) &\approx h_j(\mathbf{x}_{opt}) + \nabla\mathbf{h_j}(\mathbf{x}_{opt})^{\mathsf{T}}(\mathbf{x}-\mathbf{x}_{opt}) = 0 \Rightarrow \\ \nabla\mathbf{h_j}(\mathbf{x}_{opt})^{\mathsf{T}}(\mathbf{x}-\mathbf{x}_{opt}) &= 0 \end{aligned}$$

- Omezení ve tvaru nerovností: Uvažujeme pouze *aktivní* omezení, pro které $g_i(\mathbf{x}_{opt}) = 0$

$$\begin{aligned} g_i(\mathbf{x}) &\approx g_i(\mathbf{x}_{opt}) + \nabla\mathbf{g_i}(\mathbf{x}_{opt})^{\mathsf{T}}(\mathbf{x}-\mathbf{x}_{opt}) \leq 0 \Rightarrow \\ \nabla\mathbf{g_i}(\mathbf{x}_{opt})^{\mathsf{T}}(\mathbf{x}-\mathbf{x}_{opt}) &\leq 0 \end{aligned}$$

- Uvažujme nyní lineární kombinaci předchozích omezení (s $\mu_i > 0$)

$$\left(\sum_j \lambda_j \nabla \mathbf{h_j}(\mathbf{x}_{opt}) + \sum_{i_{akt}} \mu_i \nabla \mathbf{g_i}(\mathbf{x}_{opt})\right)^{\mathsf{T}} (\mathbf{x} - \mathbf{x}_{opt}),$$

  která je viditelně záporná.

- Najdeme-li multiplikátory

$$\lambda_{j,opt} \quad \text{a} \quad \mu_{i,opt} \geq 0$$

  splňující

$$\nabla \mathbf{f}(\mathbf{x}_{opt}) = -\left(\sum_j \lambda_{j,opt} \nabla \mathbf{h_j}(\mathbf{x}_{opt}) + \sum_{i_{akt}} \mu_{i,opt} \nabla \mathbf{g_i}(\mathbf{x}_{opt})\right)$$

  podařilo se nám splnit všechny podmínky optimality.

Shrnutí Pokud existují

- bod $\mathbf{x}_{opt}$,
- multiplikátory $\mu_{i,opt} \geq 0$ odpovídající nerovnostem a
- multiplikátory $\lambda_{i,opt}$ odpovídající rovnostem splňující podmínku

$$\nabla \mathbf{f}(\mathbf{x}_{opt}) + \left(\sum_j \lambda_{j,opt} \nabla \mathbf{h_j}(\mathbf{x}_{opt}) + \sum_{i_{akt}} \mu_{i,opt} \nabla \mathbf{g_i}(\mathbf{x}_{opt}) \right) = \mathbf{0},$$

  pak je bod $\mathbf{x}_{opt}$ hledaným optimem řešené úlohy s omezeními.

- Lagrangeovu funkci lze v tomto případě definovat jako

$$L(\mathbf{x}, \lambda, \mu, \nu) = f(\mathbf{x}) + \sum_j \lambda_j h_j + \sum_{i_{akt}} \mu_i (g_i + \nu_i)$$

  kde multiplikátory $\mu_i > 0$ a volné proměnné [slack variables] splňují $\nu_i > 0$.

- Podmínky optimality pro Lagrangeovu funkci

$$
\begin{aligned}
\frac{\partial L}{\partial \mathbf{x}}: \quad & \nabla \mathbf{f}(\mathbf{x}_{opt}) + \left(\sum_j \lambda_{j,opt} \nabla \mathbf{h_j}(\mathbf{x}) + \sum_i \mu_i \nabla \mathbf{g_i}(\mathbf{x}_{opt})\right) && = \mathbf{0}, \\
\frac{\partial L}{\partial \lambda_j}: \quad & h_j && = 0, \\
\frac{\partial L}{\partial \mu_i}: \quad & g_i + \nu_i && = 0 \\
\frac{\partial L}{\partial \nu_i}: \quad & g_i \nu_i && = 0 \\
& \mu_i && \geq 0 \\
& \nu_i && \geq 0
\end{aligned}
$$

- Zcela ekvivalentní původní formulaci
- „Vybírají“ aktivní omezení ve tvaru rovností pomocí členu $g_i \nu_i = 0$ podmínka komplementarity

Shrnutí KKT podmínek

- Základ všech metod matematického programování
- Problémem jsou „zbývající“ nerovnosti pro μ a ν
- Pro konvexní problém a „rozumná“ omezení určují řešení jednoznačně
- Nekonvexní problémy $\rightarrow$ vícenásobná řešení
- Další informace o těchto metodách viz [1]

Reference

[1] J. F. Bonnans, J. C. Gilbert, J. C. Lemarechal, and C. A. Sagastizabal, *Numerical optimization: Theoretical and practical aspects*, Springer-Verlag, Berlin, Heidelberg, 2004.

$\square$

Prosba. V případě, že v textu objevíte nějakou chybu nebo budete mít námět na jeho vylepšení, ozvěte se prosím na zemanj@cml.fsv.cvut.cz.

Verze −003