Tehnici de Optimizare

Facultatea de Matematica si Informatica

Universitatea Bucuresti

- Department Informatica-

2021

Algoritmi pentru probleme de programare convexa

Algoritmi

- Modelul convex: $\min_{x} f(x) \quad s.l.\ h(x) \leq 0$
- $h_i(x) \geq h_i(x^k) + \nabla h_i(x^k)^T (x - x^k)$

Metoda Gradient Proiectat cu Liniarizare:

$$x^{k+1} = argmin_x\, \nabla f(x^k)^T (x - x^k) + \frac{1}{2\alpha_k} \left|\left|x - x^k\right|\right|^2$$
$$s.l.\ h_i(x^k) + \nabla h_i(x^k)^T (x - x^k) \leq 0, \qquad \forall i$$

Algoritmi

Metoda gradient proiectat cu liniarizare:

$$x^{k+1} = argmin_x \, \nabla f\left(x^k\right)^T \left(x - x^k\right) + \frac{1}{2\alpha_k} \left\|x - x^k\right\|^2$$

$$s.l. \;\; h_i\left(x^k\right) + \nabla h_i\left(x^k\right)^T \left(x - x^k\right) \le 0, \qquad \forall i$$

THEOREM 1. Let the set X^* of solutions of problem (1) be nonempty, let the functions $f(x)$, $g_i(x)$ be differentiable, let their gradients satisfy a Lipschitz condition, and let Slater's condition hold. Then we can find a $\bar{\gamma} > 0$ such that for $0 < \gamma < \bar{\gamma}$ method (7) converges to a point $x^* \in X^*$. If, also, $f(x)$ is strongly convex, then $\|x^k - x^*\| \le cq^k$, $0 \le q \le 1$. □

Algoritmi

Metoda gradient proiectat cu liniarizare:

$$x^{k+1} = argmin_x \nabla f\left(x^k\right)^T\left(x - x^k\right) + \frac{1}{2\alpha_k}\left|\left|x - x^k\right|\right|^2$$
$$s.l.\ \ h_i\left(x^k\right) + \nabla h_i\left(x^k\right)^T\left(x - x^k\right) \leq 0, \qquad \forall i$$

La fiecare iteratie se rezolva:

$$\min_x \frac{1}{2}\left|\left|x - \left(x^k - \alpha_k \nabla f\left(x^k\right)\right)\right|\right|^2 = \frac{1}{2}\left|\left|x - y^k\right|\right|^2$$
$$s.l.\ Ax \leq b$$

unde $A = \begin{bmatrix} \nabla h_1\left(x^k\right)^T \\ \dots \\ \nabla h_p\left(x^k\right)^T \end{bmatrix}, b = Ax^k - h(x^k)$

Conditia Slater are loc!

MPGL

$$\min_{x} \frac{1}{2}\left|\left|x \ - y^k\right|\right|^2$$
$$s.l.\ Ax \leq b$$
$$x^{k+1} = y^k - A^T\ \tilde{u}^k$$
$$L(x,\mu) = \frac{1}{2}\left|\left|x - y^k\right|\right|^2 + \mu^T(Ax - b)$$
$$\nabla_x L(x,\mu) = 0 \Rightarrow x(\mu) = y^k - A^T\mu$$
$$\max_{\mu\geq 0} \phi(\mu) = \ (QP)$$
$$\tilde{\mu}^k \approx arg\max_{\mu\geq 0} -\frac{1}{2}||A^T\mu||^2 + \mu^T\left(Ay^k - b\right)$$

Complexitate: $O(\log(1/\epsilon))$ iteratii necesare pentru a atinge: $\left|\left|x^k - x^*\right|\right| \leq \epsilon$

Metoda Gradient Dual

Modelul convex: $\min_x f(x) \quad s.l.\ h(x) \leq 0\,, Ax = b$

Problema duala:

$$\max_{\boldsymbol{\lambda},\boldsymbol{\mu}\geq \mathbf{0}} \boldsymbol{\phi}(\boldsymbol{\lambda},\boldsymbol{\mu}) = -\boldsymbol{min}_{\boldsymbol{\lambda},\boldsymbol{\mu}\geq \mathbf{0}} - \boldsymbol{\phi}(\boldsymbol{\lambda},\boldsymbol{\mu})$$

- Pp: Problema convexa, conditia Slater are loc!
- Constrangeri doar pe μ!
- Daca functia obiectiv primala f este tare convexa (σ_f) atunci ϕ este diferentiabila cu gradient Lipschitz;

$$\begin{bmatrix} \lambda^{k+1} \\ \mu^{k+1} \end{bmatrix} = \boldsymbol{\pi}_{\boldsymbol{Q}}\left(\begin{bmatrix} \lambda^{k} \\ \mu^{k} \end{bmatrix} + \boldsymbol{\alpha}_{\boldsymbol{k}} \boldsymbol{\nabla\phi}\left(\begin{bmatrix} \lambda^{k} \\ \mu^{k} \end{bmatrix}\right)\right)$$

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T g(x) + \mu^T h(x) \rightarrow x(\lambda, \mu)\ (solutie)$$

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T (Ax - b) + \mu^T h(x) \rightarrow x(\lambda, \mu)\ (solutie)$$
$$\nabla_\lambda \phi(\lambda, \mu) = Ax(\lambda, \mu) - b$$

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$

$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T (Ax - b) + \mu^T h(x) \rightarrow x(\lambda, \mu)\ (solutie)$$

$$\nabla_\lambda \phi(\lambda, \mu) = Ax(\lambda, \mu) - b \in R^m$$

$$\nabla_\mu \phi(\lambda, \mu) = h\left(x(\lambda, \mu)\right) \in R^p$$

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T g(x) + \mu^T h(x) \rightarrow x(\lambda, \mu)\ (solutie)$$
$$\nabla_\lambda \phi(\lambda, \mu) = g\left(x(\lambda, \mu)\right) \qquad \nabla_\mu \phi(\lambda, \mu) = h(x(\lambda, \mu))$$

Pentru calculul $\nabla\phi$ este necesara solutia $x(\lambda, \mu)$ a problemei $\min\limits_{\mathrm{x}} L(x, \lambda, \mu)$!

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T g(x) + \mu^T h(x) \rightarrow x(\lambda, \mu)\ (solutie)$$
$$\nabla_\lambda \phi(\lambda, \mu) = g\left(x(\lambda, \mu)\right) \qquad \nabla_\mu \phi(\lambda, \mu) = h(x(\lambda, \mu))$$

Pentru calculul $\nabla\phi$ este necesara solutia **aproximativa** $\tilde{x}\ (\lambda, \mu)$ a problemei $\min_{\mathrm{x}} L(x, \lambda, \mu)$!

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$

$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T (Ax - b) + \mu^T h(x) \rightarrow x(\lambda, \mu)\ (solutie)$$

$$\nabla_\lambda \phi(\lambda, \mu) = g\left(x(\lambda, \mu)\right) \qquad \nabla_\mu \phi(\lambda, \mu) = h(x(\lambda, \mu))$$

Pentru calculul $\nabla\phi$ este necesara solutia **aproximativa** $\tilde{x}\ (\lambda, \mu)$ a problemei $\min_{\mathrm{x}} L(x, \lambda, \mu)$!

Daca f, h sunt diferentiabile/liniare atunci $\tilde{x}(\lambda, \mu)$ se obtine folosind orice algoritm iterative pentru probleme fara constrangeri.

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

$$\phi(\lambda, \mu) = \min_x f(x) + \lambda^T g(x) + \mu^T h(x) \to x(\lambda, \mu)\ (solutie)$$

Oprire: Se opreste MGPD cu acelasi criteriu ca MGP, iar iteratia finala $(\lambda^K, \mu^K) final$ se foloseste pentru $x(\lambda^K, \mu^K) \approx x^*$

Metoda Gradient Dual

Metoda Gradient Proiectat Dual:

$$\lambda^{k+1} = \lambda^k + \alpha_k \nabla_\lambda \phi\left(\lambda^k, \mu^k\right)$$
$$\mu^{k+1} = \pi_+\left(\mu^k + \alpha_k \nabla_\mu \phi(\lambda^k, \mu^k)\right)$$

Teorema. Presupunem ca conditia Slater are loc in problema primala si f tare convexa. Atunci ϕ are gradient Lipschitz si MGPD converge:

$$\phi^* - \phi\left(\lambda^k, \mu^k\right) \leq O\left(\frac{1}{k}\right)$$

Weisfeld

$$\min_x \sum_i \omega_i ||x - a_i||^2$$

Weisfeld

$$\min_x \sum_i \omega_i ||x - a_i||^2$$

Functia obiectiv este convexa:

$$\nabla f(x^*) = 0$$

$$\sum_i \omega_i (x^* - a_i) = 0$$

Weisfeld

$$\min_x \sum_i \omega_i \left| |x - a_i| \right|^2$$

Functia obiectiv este convexa:

$$\nabla f(x^*) = 0$$

$$\sum_i \omega_i (x^* - a_i) = 0$$

$$x^* \sum_i \omega_i = \sum_i \omega_i a_i$$

Weisfeld

$$\min_x \sum_i \omega_i ||x - a_i||^2$$

Functia obiectiv este convexa:

$$\nabla f(x^*) = 0$$

$$\sum_i \omega_i (x^* - a_i) = 0$$

$$x^* \sum_i \omega_i = \sum_i \omega_i a_i$$

$$x^* = \sum_i \frac{\omega_i a_i}{\sum_i \omega_i}$$

- $\nabla|x| = \frac{x}{|x|} = \mathrm{sgn}(x)\,, pt\ x \neq 0 \Rightarrow \nabla|x-a| = \frac{x-a}{|x-a|}$
- $\nabla|0| = [-1{,}1]; s \in [-1{,}1] = \partial f(0)$ se numeste subgradient
- $|\nabla f(x) - \nabla f(y)| \leq L|x-y| \rightarrow |\,\mathrm{sgn}(x) - \mathrm{sgn}(y)\,| < L|x-y|$
- $x < 0, y < 0 \Rightarrow 2 \leq L|x-y| \rightarrow 0$

- $g(x) = \left||x-a_i|\right|, g\ diferentiabila\ cu\ exceptia\ x^* = a_i$ (nu are grad. Lipschitz)
- $\nabla \mathrm{g}(\mathrm{x}) = \frac{\mathrm{x}-\mathrm{a_i}}{\left||\mathrm{x}-\mathrm{a_i}|\right|}, \mathrm{x} \neq a_i$
- $\nabla g(a_i) = \partial g(a_i) = \{x: \left||x-a_i|\right| \leq 1\}$

- $g(x) = \left|\left|x - a_i\right|\right|, g\ diferentiabila\ cu\ exceptia\ x^* = a_i$ (nu are grad. Lipschitz)
- $\nabla \mathrm{g}(\mathrm{x}) = \frac{\mathrm{x}-\mathrm{a_i}}{\left|\left|\mathrm{x}-\mathrm{a_i}\right|\right|}, \mathrm{x} \neq a_i$
- $\nabla g(a_i) = \partial g(a_i) = \{x: \left|\left|x - a_i\right|\right| \leq 1\}$
- $\nabla f(x^*) = \frac{\sum_i \omega_i (x^* - a_i)}{\left|\left|x^* - a_i\right|\right|} = 0 \Rightarrow \omega_i v = -\frac{\sum_{i \neq j} \omega_i (x^* - a_i)}{\left|\left|x^* - a_i\right|\right|}, unde \left|\left|v\right|\right| \leq 1$
- $\left|\left|\frac{1}{\omega_i} \frac{\sum_{i \neq j} \omega_i (x^* - a_i)}{\left|\left|x^* - a_i\right|\right|}\right|\right| \leq 1$