An optimal algorithm with Barzilai-Borwein steplength and superrelaxation for QPQC problem

[Panm16]

Lukáš Pospíšil

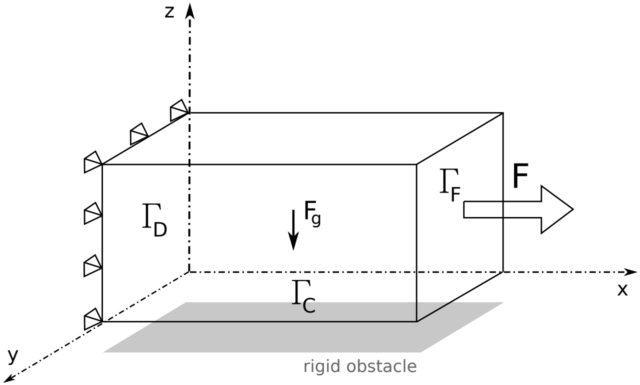
Department of Applied Mathematics,
VŠB - Technical University of Ostrava

5.6.2012

Osnova prezentace

- Motivace (kontaktní úloha s daným třením)
- QPQC
- stávající algoritmy (MPGP, SPG)
- modifikace (MPGP-BB)
- numerické testy
- několik poznámek ke konvergenci
- závěr

Kontaktní úloha s daným třením



Primární úloha

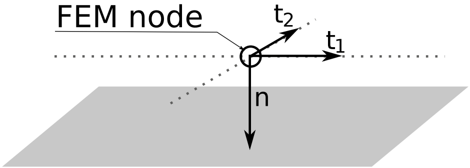
Nalezněte

$$\bar{\mathbf{u}} := \arg\min_{\mathbf{u}\in\Omega_C} f(\mathbf{u}) + j_h(\mathbf{u})$$

kde

- $\Omega_C := \{\mathbf{u} \in \Gamma_C : u_z \geq -d_C\}$,
- $f(\mathbf{u}) := 1/2\langle \mathbf{K}\mathbf{u}, \mathbf{u}\rangle - \langle \mathbf{f}, \mathbf{u}\rangle, f : \mathbb{R}^n \to \mathbb{R}$,
- $j_h(\mathbf{u}) := \sum_{i=1}^{m_c} \psi_i \|\mathbf{T}_i\mathbf{u}\|, j_h : \mathbb{R}^n \to \mathbb{R}$.

Tření a tečné vektory

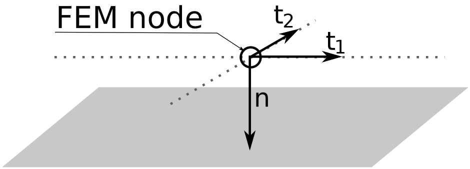


$$\mathbf{T}_i := \begin{bmatrix} 0 & 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{T} := \left[\mathbf{T}_1^T, \dots, \mathbf{T}_{m_c}^T\right]^T$$

$$j_h(\mathbf{u}) := \sum_{i=1}^{m_c} \psi_i \|\mathbf{T}_i\mathbf{u}\| = \sum_{i=1}^{m_c} \max_{\|\tau_i\| \le \psi_i} \boldsymbol{\tau}_i^T \mathbf{T}_i \mathbf{u}$$

Podmínky nepronikání a normálové vektory



$$r_i^{\mathbf{B}} := \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 & -1 & \dots & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{c}_i := d_C$$

$$\mathbf{Bu} \leq \mathbf{c}$$

Duální úloha

Z primární úlohy

$$\min_{\mathbf{u}\in\Omega_C}\left(f(\mathbf{u})+\sum_{i=1}^{m_c}\max_{\|\tau_i\|\le\psi_i}\boldsymbol{\tau}_i^T\mathbf{T}_i\mathbf{u}\right)$$

vytvořením nových Lagrangeových multiplikátorů

$$\min_{\mathbf{u}\in\mathbb{R}^n}\sup_{[\tau,\boldsymbol{\lambda}_C]\in\Lambda} f(\mathbf{u})+\boldsymbol{\tau}^T\mathbf{T}\mathbf{u}+\boldsymbol{\lambda}_C^T(\mathbf{B}\mathbf{u}-\mathbf{c}),$$

kde

$$\Lambda:=\{[\boldsymbol{\tau},\boldsymbol{\lambda}_C]\in\mathbb{R}^{3m_c}:\sqrt{\tau_{2i-1}^2+\tau_{2i}^2}\le\psi_i, i=1,\dots,m_c,\boldsymbol{\lambda}_C\ge\mathbf{o}\}.$$

Dále pak

$$\min_{\mathbf{u}\in\mathbb{R}^n}\sup_{\boldsymbol{\lambda}\in\Lambda} f(\mathbf{u})+\boldsymbol{\lambda}^T(\tilde{\mathbf{B}}\mathbf{u}-\tilde{\mathbf{c}})=:\min_{\mathbf{u}\in\mathbb{R}^n}\sup_{\boldsymbol{\lambda}\in\Lambda} L(\mathbf{u},\boldsymbol{\lambda}),$$

kde

$$\boldsymbol{\lambda}:=\begin{bmatrix}\boldsymbol{\tau}\\ \boldsymbol{\lambda}_C\end{bmatrix},\quad \tilde{\mathbf{B}}:=\begin{bmatrix}\mathbf{T}\\ \mathbf{B}\end{bmatrix},\quad \tilde{\mathbf{c}}:=\begin{bmatrix}\mathbf{o}\\ \mathbf{c}\end{bmatrix}.$$

Duální úloha

$$\min_{\mathbf{u}\in\mathbb{R}^n} \sup_{\boldsymbol{\lambda}\in\Lambda} L(\mathbf{u},\boldsymbol{\lambda}) = \max_{\boldsymbol{\lambda}\in\Lambda} \min_{\mathbf{u}\in\mathbb{R}^n} L(\mathbf{u},\boldsymbol{\lambda}).$$

Z první KKT podmínky

$$\begin{aligned} \mathbf{K}\mathbf{u} - \mathbf{f} + \tilde{\mathbf{B}}^T\boldsymbol{\lambda} &= \mathbf{o} \\ \mathbf{u} &= \mathbf{K}^{-1}\left(\mathbf{f} - \tilde{\mathbf{B}}^T\boldsymbol{\lambda}\right). \end{aligned}$$

Dosazením

$$L(\mathbf{u},\boldsymbol{\lambda}) = -\frac{1}{2}\boldsymbol{\lambda}^T\tilde{\mathbf{B}}\mathbf{K}^{-1}\tilde{\mathbf{B}}^T\boldsymbol{\lambda} + \boldsymbol{\lambda}^T(\tilde{\mathbf{B}}\mathbf{K}^{-1}\mathbf{f} - \tilde{\mathbf{c}}) - \frac{1}{2}\mathbf{f}^T\mathbf{K}^{-1}\mathbf{f}.$$

Duální úloha má tvar

$$\bar{\boldsymbol{\lambda}} = \arg\min_{\lambda\in\Lambda} \Theta(\boldsymbol{\lambda}), \quad \Theta(\boldsymbol{\lambda}) := \frac{1}{2}\boldsymbol{\lambda}^T\mathbf{F}\boldsymbol{\lambda} - \boldsymbol{\lambda}^T\mathbf{d},$$

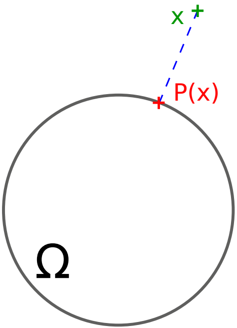
kde

$$\mathbf{F} := \tilde{\mathbf{B}}\mathbf{K}^{-1}\tilde{\mathbf{B}}^T, \quad \mathbf{d} := \tilde{\mathbf{B}}\mathbf{K}^{-1}\mathbf{f} - \tilde{\mathbf{c}}.$$

QPQC

$$\Omega := \{\mathbf{x} \in \mathbb{R}^n : \sqrt{x_{2i-1}^2 + x_{2i}^2} \leq r_i, \text{ for all } i = 1, \dots, m\}.$$

$$P_\Omega(\mathbf{x}) := \begin{bmatrix} \frac{r_1}{\|[x_1, x_2]\|}[x_1, x_2]^T \\ \vdots \\ \frac{r_i}{\|[x_{2i-1}, x_{2i}]\|}[x_{2i-1}, x_{2i}]^T \\ \vdots \\ \frac{r_m}{\|[x_{2m-1}, x_{2m}]\|}[x_{2m-1}, x_{2m}]^T \end{bmatrix}$$



Modified proportioning with gradient projections (MPGP)

Algoritmus 1 MPGP

```
Choose $x_0 \in \Omega, \alpha \in (0, 2\|A\|^{-1}), \delta \in (0, 1/2\rangle$
$k := 0$
while $\|\tilde{g}_k\| \geq \epsilon\|b\|$ do
    if $2\delta g_k^T g_k^P \leq \|\varphi(x_k)\|^2$ then
        CG step or CG halfstep.
        make CG step to solve problem on free set.
        if this step means leaving $\Omega$, do only a half-step and restart CG.
        $k := k + 1$
    else
        Projection step.
        $x_{k+1} := P(x_k - \alpha g_k)$
        restart CG
        $k := k + 1$
    end if
end while
```

Spectral projected gradient method (SPG)

Algoritmus 2 SGP

Choose $\boldsymbol{x}_0 \in \Omega,\ 0 < \alpha_{min} \leq \alpha_0 \leq \alpha_{max},\ 0 < \sigma_1 < \sigma_2 < 1,\ \gamma \in (0,1)$

$k := 0$

while $\|P(\boldsymbol{x}_k - \alpha_k \boldsymbol{g}_k) - \boldsymbol{x}_k\| > \epsilon$ **do**

  $\boldsymbol{d}_k := P(\boldsymbol{x}_k - \alpha_k \boldsymbol{g}_k) - \boldsymbol{x}_k$

  Find $\lambda \in \langle \sigma_1, \sigma_2 \rangle$ such that

$$f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k) \leq f(\boldsymbol{x}_k) + \gamma \lambda \langle \boldsymbol{d}_k, \boldsymbol{g}_k \rangle$$

  $\boldsymbol{x}_{k+1} := \boldsymbol{x}_k + \lambda \boldsymbol{d}_k$

  $\boldsymbol{s}_k := \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$

  $\boldsymbol{y}_k := \boldsymbol{g}_{k+1} - \boldsymbol{g}_k$

  $\alpha_{k+1} := \min\{\alpha_{max}, \max\{\alpha_{min}, \frac{\langle \boldsymbol{s}_k, \boldsymbol{s}_k \rangle}{\langle \boldsymbol{y}_k, \boldsymbol{s}_k \rangle}\}\}$

  $k := k + 1$

end while

Modified proportioning with Barzilai-Borwein gradient projections (MPGP-BB)

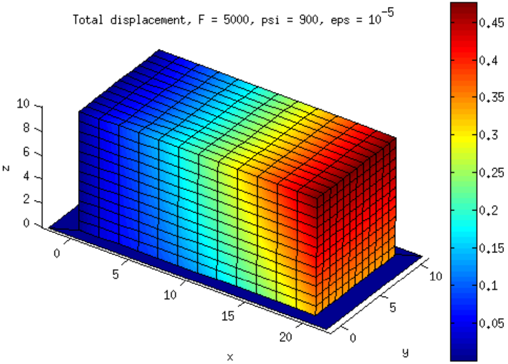
- měnit délku projektovaného gradientu dle SPG *bez* dodatečného GLL (uvolnění vazby, $f(\mathbf{x})$)

Algoritmus 3 MPGP-BB

Choose $\boldsymbol{x}_0 \in \Omega, \alpha \in (0, 2\|A\|^{-1}), \delta \in (0, 1/2\rangle$
$\alpha_{bb} := \alpha$
$k := 0$
while $\|\tilde{\boldsymbol{g}}_k\| \geq \epsilon\|b\|$ **do**
  if $2\delta \boldsymbol{g}_k^T \boldsymbol{g}_k^P \leq \|\varphi(\boldsymbol{x}_k)\|^2$ **then**
    CG step or CG halfstep.
    make CG step to solve problem on free set.
    if this step means leaving Ω, do only a half-step and restart CG.
    $k := k + 1$
  else
    Projection step.
    $\boldsymbol{x}_{k+1} := P(\boldsymbol{x}^k - \alpha_{bb}\boldsymbol{g}_k)$
    $\boldsymbol{s} := \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$
    $\alpha_{bb} := \boldsymbol{s}^T\boldsymbol{s}/\boldsymbol{s}^T\boldsymbol{A}\boldsymbol{s}$
    restart CG
    $k := k + 1$
  end if
end while

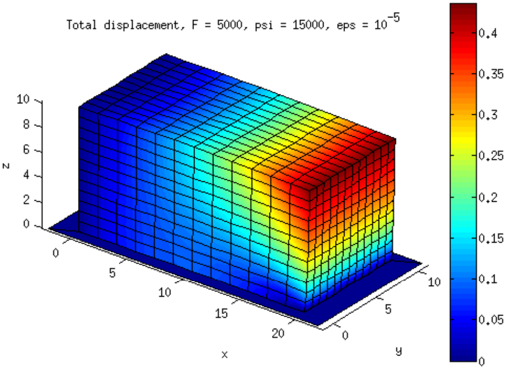
Numerické testy

Pro implementaci využit Matlab + MatSOL.
Konkrétní hodnoty: ocel, $\psi = 900$, $F = 5000$, $\epsilon = 10^{-5}$
primární proměnné: 6591, duální proměnné: 507
aktivní vazby: 93/169 Q, 39/169 L



Numerické testy

Konkrétní hodnoty: ocel, $\psi = 15000, F = 5000, \epsilon = 10^{-5}$
primární proměnné: 6591, duální proměnné: 507
aktivní vazby: 0/169 Q, 35/169 L



Numerické testy

F = 5000, ψ = 900, $\epsilon = 10^{-4}$												
N	**primal**	**dual**	**SPG**	**MPGP**				**MPGP-BB**				**active**
			it	it	cg	half	proj	it	cg	half	proj	Q,L
4	375	75	36	5176	0	1	5175	41	0	1	40	11/25, 10/25
6	1029	147	45	2746	0	1	2745	57	0	1	56	27/49, 21/49
8	2187	243	27	1236	0	1	1235	51	0	1	50	44/81, 27/81
10	3993	363	33	661	0	1	660	40	0	1	39	76/121, 33/121

F = 5000, ψ = 5000, $\epsilon = 10^{-4}$												
N	**primal**	**dual**	**SPG**	**MPGP**				**MPGP-BB**				**active**
			it	it	cg	half	proj	it	cg	half	proj	Q,L
4	375	75	1128	3191	177	15	2999	216	15	25	176	3/25, 10/25
6	1029	147	479	2005	25	18	1962	139	8	53	78	5/49, 14/49
8	2187	243	341	1860	5	22	1833	125	1	45	79	6/81, 23/81
10	3993	363	199	1078	2	30	1046	103	2	51	50	5/121, 31/121

F = 5000, ψ = 15000, $\epsilon = 10^{-4}$												
N	**primal**	**dual**	**SPG**	**MPGP**				**MPGP-BB**				**active**
			it	it	cg	half	proj	it	cg	half	proj	Q,L
4	375	75	1566	43	33	9	1	43	33	9	1	0/25, 10/25
6	1029	147	923	48	29	18	1	48	29	18	1	0/49, 14/49
8	2187	243	588	53	24	28	1	53	24	28	1	0/81, 18/81
10	3993	363	1020	101	40	46	15	73	27	40	6	0/121, 27/121

Několik poznámek ke konvergenci

- **MPGP** (Dostál, Schöberl, Kozubek) projekce s konstantní délkou kroku $\alpha \in (0, 2/\|A\|)$ zaručuje pokles cenové funkce v každé iteraci

$$\mathbf{x}_{k+1} := P(\mathbf{x}_k - \alpha \mathbf{g}_k)$$
$$f(\mathbf{x}_{k+1}) - f(\bar{\mathbf{x}}) \leq (1 - \alpha \lambda^{\mathbf{A}}_{\max})(f(\mathbf{x}_k) - f(\bar{\mathbf{x}}))$$

- **SPG** (Martínez, Birgin, Raydan) konvergence je zaručena splněním Armijo podmínky v každé iteraci

$$\mathbf{d}_k := P(\mathbf{x}_k - \alpha_k \mathbf{g}_k) - \mathbf{x}_k$$
$$\text{find} \quad \lambda : \quad f(\mathbf{x}_k + \lambda \mathbf{d}_k) \leq f(\mathbf{x}_k) + \gamma \lambda \langle \mathbf{d}_k, \mathbf{g}_k \rangle$$
$$\mathbf{x}_{k+1} := \mathbf{x}_k + \lambda \mathbf{d}_k$$

Několik poznámek ke konvergenci

- **MPGP-BB**
  - délka kroku není konstatní a numerické testy naznačují, že pokles cenové funkce není monotónní (BB není monotónní ani v neprojektované verzi)
  - původní Raydanův důkaz neprojektovaného BB je založen na konvergenci souřadnic chyby $\mathbf{x}_k - \hat{\mathbf{x}}$ v bázi vlastních vektorů matice $\mathbf{A}$
  - Armijo podmínka není splněna v každé iteraci

Několik poznámek ke konvergenci

Pro QP na konvexních množinách se mi *zatím* podařilo odvodit

$$f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_k) + \gamma_k \langle \mathbf{x}_{k+1} - \mathbf{x}_k, \mathbf{g}_k \rangle$$

kde

$$\gamma_k := \left(1 - \frac{\alpha_k}{2\alpha_{k+1}}\right) .$$

Děkuji za pozornost

Birgin, E.G., Martínez, J.M., and Raydan, M.: *Nonmonotone spectral projected gradient methods on convex sets*, SIAM J. on Optimization 10, 1196–1211., 2000.

Dostál, Z.: *Optimal Quadratic Programming Algorithms, with Applications to Variational Inequalities*, SOIA, first ed., vol. 23, Springer, US, New York, 2009.

Dostál, Z., Kozubek, T.: *An optimal algorithm and superrelaxation for minimization of a quadratic function subject to separable convex constraints with applications*, SOIA, first ed., vol. 23, Springer, US, New York, 2009.

Dostál, Z., Kozubek, T., Markopoulos, A., Brzobohatý, T., Vondrák, V., Horyl, P.: *A theoretically supported scalable TFETI algorithm for the solution of multibody 3D contact problems with friction*, Computer methods in applied mechanics and engineering. 2012, vol. 205-208, p. 110-120, 2012.

This work was supported by grant GD103/09/H078 of Grant Agency of the Czech Republic.

The author would like to express his special thanks to
Prof. RNDr. Zdeněk Dostál, DSc., Ing. Alexandros Markopoulos, Ph.D. and Ing. Tomáš Brzobohatý, Ph.D.