Integration of bound constraints in a recursive multilevel trust-region algorithm

M. Mouffe
CERFACS, Toulouse

10th October 2007

joint work with

S. Gratton
CERFACS, Toulouse

Ph. L. Toint
FUNDP, Namur

Outline

Trust-region mechanism

[A. Conn, N. Gould and Ph. L. Toint, 2000]

- Define a model m_k of the objective function f
- Define a trust region where the model is supposed to represent well the objective function
- Compute a step (TR subproblem)
  - inside the TR (in Euclidean or infinite norm)
  - that sufficiently reduces m_k
- Step acceptance and TR radius Δ update related to the ratio

$$\frac{f(x_{k+1}) - f(x_k)}{m_k(x_{k+1}) - m_k(x_k)}$$

  - Refuse the step and shrink the TR when the ratio is smaller than a constant
  - Accept the step and possibly enlarge the TR when the ratio is large enough

Why multigrid

- Solution based on discretization :
  High accuracy $\Rightarrow$ computational cost
- Use of coarse grids :
  1. find a good starting point
  2. solve a subproblem (e.g. the TR subproblem)
- Well-known for solving SPD linear systems resulting of the discretization of a continuous problem
  [W. Briggs, V.E. Henson and S. McCormick, 2000]
- Nonlinear systems
  [W. Hackbuch and A. Reusken, 1989]

Outline

Introduction

▶ Multilevel trust-region algorithm

Bound constraints

Conclusion and perspectives

Multigrid

- Suppose you have a set of discretizations of the objective function f :

$$\{f_i\}_{i=0}^{r}$$

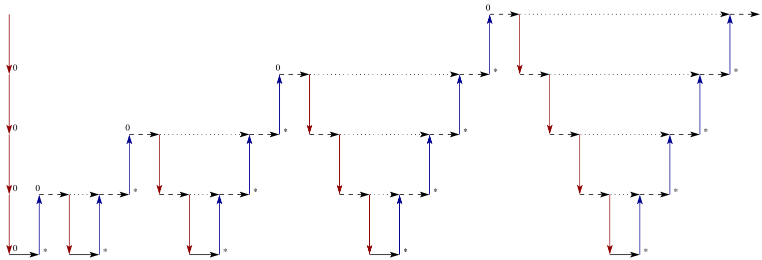
with $f_r = f$.

- Define transfert operators :

$$\begin{array}{llcll} R_i : & \mathbb{R}^{n_i} & \longrightarrow & \mathbb{R}^{n_{i-1}} & \text{Restriction} \\ P_i : & \mathbb{R}^{n_{i-1}} & \longrightarrow & \mathbb{R}^{n_i} & \text{Prolongation} \end{array}$$

- Use coarser levels of discretization to :
  1. Find a good starting point
  2. Solve the TR subproblem

Full multilevel scheme (FMG) : Combination of mesh refinement and V-cycles



Outline

Theoretical point of view

Starting from RMTR :

- Adapt the criticality measure
- Ensure sufficient decrease at smoothing iterates
- Ensure step properties :
  1. Trust region
  2. Feasibility

  for iterates computed on a coarser level

Convergence

First adapt like the classical TR method

- Use of an infinity norm trust region
- Criticality measure :

$$||\nabla_x f_{i,k}(x_{i,k})|| \quad \Rightarrow \quad \chi_{i,k} = \left| \min_{\substack{x_{i,k}+d \in \mathcal{C} \\ ||d||_\infty \leq 1}} \nabla_x f_{i,k}(x_{i,k})^T d \right|$$

[A. Conn, N. Gould and Ph. L. Toint, 2000]

- Sufficient decrease :

$$m_{i,k}(x_{i,k+1}) - m_{i,k}(x_{i,k}) \geq \kappa \chi_{i,k} \min \left[\frac{\chi_{i,k}}{\beta_{i,k}}, \Delta_{i,k}, 1 \right]$$

- Stopping criterion : $\chi_{i,k} \leq \varepsilon$ (i fine)
- Descent condition : $\chi_{i-1} > \kappa_g \chi_i$ ($i-1$ coarse)

Zoom on the coarse step computation

It is sufficient that $s_{i,k} = P s_{i-1,*}$

- lays inside the TR (bounded violation)
- is feasible with respect to bound constraints

[S. Gratton, M. Mouffe, Ph. L. Toint, M. Weber Mendonça, 2007]

Step inside the TR

- Coarse trust region definition :
  1. Use R on the upper TR
  2. Local independant TR
  3. Intersect to define the "working" TR
- Consequence :
  1. and 3. $\Rightarrow$ The prolongation of the coarse step is inside a multiplicative constant of the TR

Possible treatments of the coarse bounds

1. Linearly constrained problem
2. Simple restriction of the bounds
3. Adapt E. Gelman and J. Mandel's definition
4. Add a security layer to 3.
5. ...

1. Linearly constrained problem

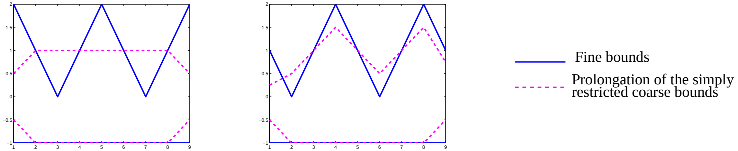
$$l \leq P_r \ldots P_i s_{i,k} \leq u$$

- Risk : Expensive to solve even on a coarser level
- Recursive idea is lost
- Maintain bound constrained nature of the problem at all levels

2. Simple restriction of the fine bounds

- Definition : $l_{i-1} = Rl_i$ and $u_{i-1} = Ru_i$
- Possibility to have infeasible iterates :

$$PRl \not\geq l \text{ or } PRu \not\leq u$$



2. Simple restriction of the fine bounds

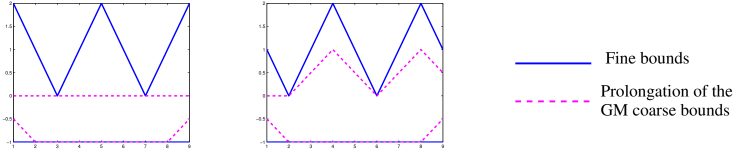
- Solution : projection of the infeasible iterate on the fine bounds
- Need to check sufficient decrease
- Risk : many rejections of coarse steps
- Strategies to limit the length of the coarse step

3. Adaptation of Gelman & Mandel's definition

- Definition :

$$\begin{cases} [l_{r-1}]_j = [R_r x_{r,k}]_j + \max\limits_{\substack{t=1,\dots,n_r \\ [P_r]_{tj} > 0}} [l - x_{r,k}]_t \\ [u_{r-1}]_j = [R_r x_{r,k}]_j + \min\limits_{\substack{t=1,\dots,n_r \\ [P_r]_{tj} > 0}} [u - x_{r,k}]_t \end{cases}$$

[E. Gelman and J. Mandel, 1990]



- Generalization to P negative and $||P||_\infty > 1$

3. Adaptation of Gelman & Mandel's definition

- NO infeasible iterates
  $\Rightarrow$ Convergence
- BUT avoid too restrictive steps
  $\Rightarrow$ Authorize controled infeasibility
- Define a security layer around the GM coarse set of constraints
- BUT checking the sufficient decrease condition remains costly

Active constraints identification

- Constraint activation at the uppest level
  Theory OK
- Try to adapt other identification techniques :
  - LBFGSB
    [R. H. Byrd, P. Lu, J. Nocedal and C. Y. Zhu, 1995]
  - ASA
    [W. W. Hager and H. Zhang, 2006]
  - ASTRAL
    [L. Xu and J. V. Burke, 2007]
  - Expansion step
    [M. Domoradova and Z. Dostal, 2007]

Minimal surface problem

Solve

$$\min_u \int_a^b \int_a^b \sqrt{1 + u_x^2 + u_y^2}\, dxdy$$

Boundary conditions :

$$\begin{cases} r_0 = r_2 & = & x(1-x) \\ r_1 = r_3 & = & 0 \end{cases}$$

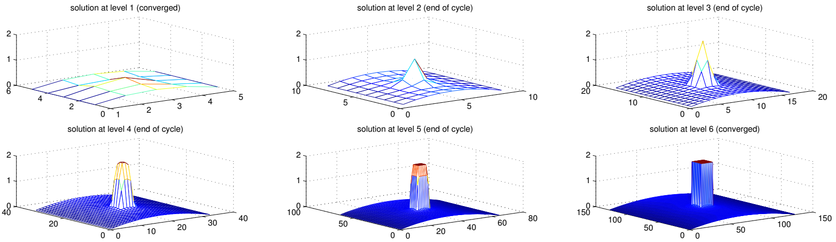
Constraints :

$$\begin{cases} lb & = & 2 \quad \text{if } \frac{4(b-a)}{9} \leq x \leq \frac{5(b-a)}{9} \\ & & 0 \quad \text{otherwise} \\ ub & = & \infty \end{cases}$$

Variables : $\approx 10^6$

Results

	Mesh Refinement	$RMTR_\infty$
Nb iter	4009	125
Nb eval f	1551	146
Nb eval g	1547	142
Nb eval H	40	24



Outline

Conclusion

- RMTR has become a bound constrained solver
- Convergence theory
- Good results with Matlab code
- Adaptation to Fortran (with D. Tomanos, FUNDP, Namur, Belgium)
- Looking forward to integrating usual bound management techniques in RMTR

Thank you !