A comparative analysis of various Trust Region subproblems

Saurav Shetty
Rahul Hegde
Rean Fernandes

Albert-Ludwigs-Universität Freiburg

UNI FREIBURG

Aim

Implementation of Trust region subproblem solution variants to arrive at a minima of a given function.

- Cauchy point method
- Dogleg method
- Conjugate Gradient Steihaug method

Comparative analysis of the trust region subproblem variants.

- Scaling factor
- Run time and iterations
- Convergence rates

Objective functions used:

- Rosenbrock
- Booth
- Ackley

Standard Algorithm

Set the starting point at x_0, set the iteration number $k = 1$

for $k = 1, 2...$

Get the improving step by solving trust-region sub-problem ()

Evaluate ρ_k from equation()

if $\rho_k < \eta_2$

$\Delta_{k+1} = t_1 \Delta_k$

else

if $\rho_k > \eta_3$ and $p_k = ||\Delta_k||$ (full step and model is a good approximation)

$\Delta_{k+1} = min(t_2 \Delta_k, \Delta_M)$

else

$\Delta_{k+1} = \Delta_k$

if $\rho_k > \eta_1$

$x_{k+1} = x_k + p_k$

else

$x_{k+1} = x_k$ (the model is not a good approximation and need to solve another trust-region subproblem within a smaller trust-region)

end >

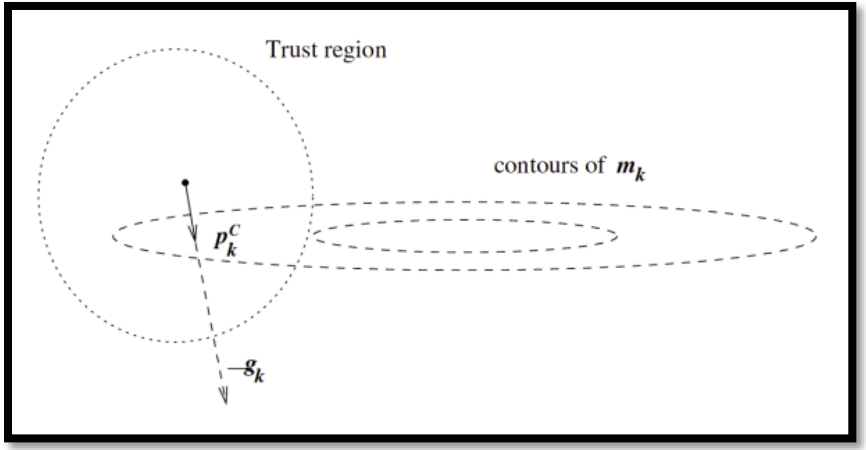
$$\boxed{\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}}$$

Cauchy point method

$$p_k{}^C = -\tau_k \frac{\Delta_k}{||g_k||} g_k$$

if $g_k{}^T B_k g_k <= 0 \ \ \tau_k = 1$

otherwise $\tau_k = min\ (||g_k||^3/(\Delta_k g_k{}^T B_k g_k), 1)$



Limitations:

- Though Cauchy point is cheap to implement, it performs poorly in some cases.
- Various kinds of improvements are based on including the curvature information from B_k.

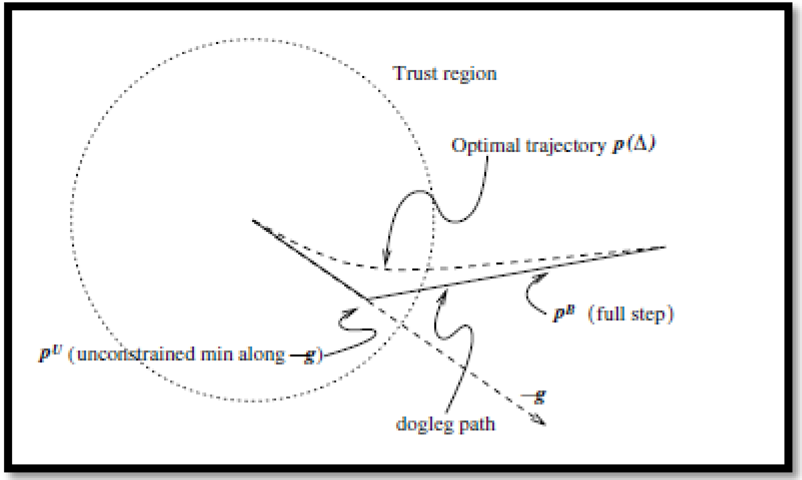
Dogleg method

If B_k is positive definite (we can use quasi-Newton Hessian approximation &updating to guarantee), then a V-shaped trajectory can be determined by

if $0 <= \tau <= 1, \quad p(\tau) = \tau p^U$

if $1 <= \tau <= 2, \quad p(\tau) = \tau p^U + (\tau - 1)(p^B - p^U)$

where $p^U = -\dfrac{g^T g}{g^T B g} g$ is the steepest descent direction and p^B is the optimal solution of the quadratic model $m_k(p)$. Therefore, a further improvement could be achieved compared to using only Cauchy point calculation method in one iteration. (Note that hessian or approximate hessian will be evaluated in dogleg method)



Conjugate Gradient Steihaug method

Given tolerance $\epsilon_k > 0$;

Set $z_0 = 0, r_0 = \nabla f_k, d_0 = -r_0 = -\nabla f_k$

if $||r_0|| < \epsilon_k$

return $p_k = z_0 = 0$;

for $j = 0, 1, 2, \ldots$

if $d_j{}^T B_k d_j <= 0$

Find τ such that $p_k = z_j + \tau d_j$ minimizes $m_k(p_k)$

and satisfies $||p_k|| = \Delta_k$;

return p_k ;

Set $\alpha_j = r_j{}^T r_j / d_j{}^T B_k d_j$;

Set $z_{j+1} = z_j + \alpha_j d_j$;

if $||z_{j+1}|| >= \Delta_k$

Find $\tau >= 0$ such that $p_k = z_j + \tau d_j$ satisfies $||p_k|| = \Delta_k$;

return p_k ;

Set $r_{j+1} = r_j + \alpha_j B_k d_j$;

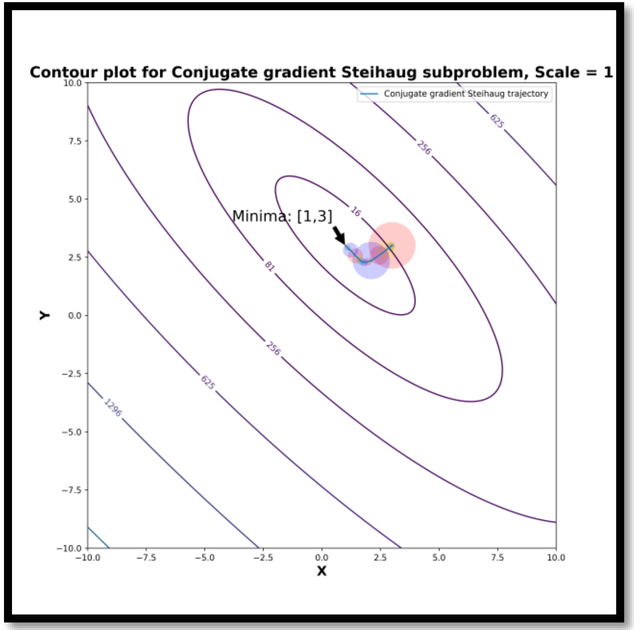
if $||r_{j+1}|| < \epsilon_k$

return $p_k = z_{j+1}$;

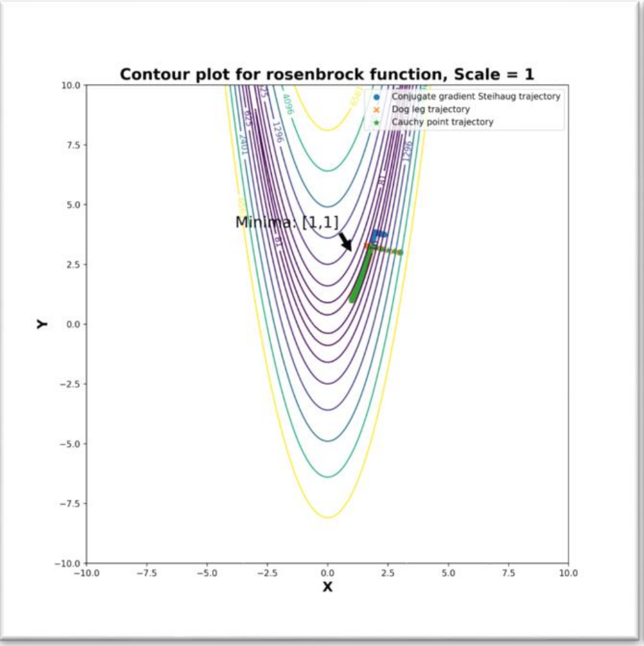
Set $\beta_{j+1} = \dfrac{r_{j+1}{}^T r_{j+1}}{r_j{}^T r_j}$;

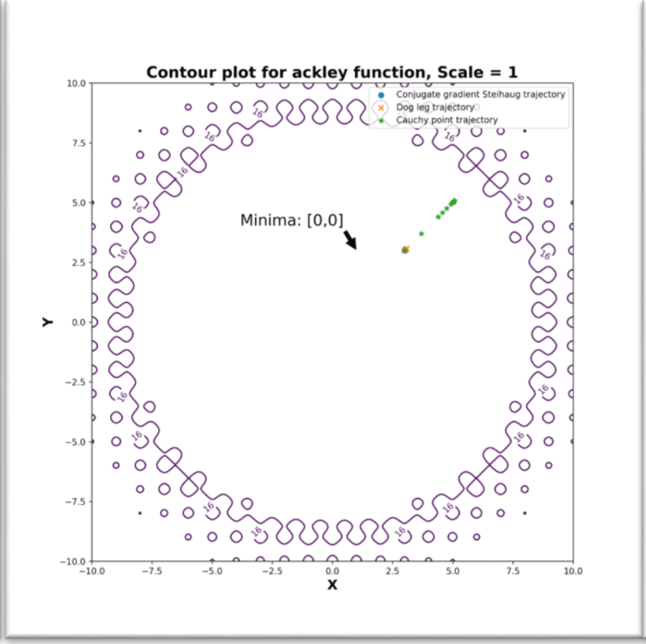
Set $d_{j+1} = -r_{j+1} + \beta_{j+1} d_j$

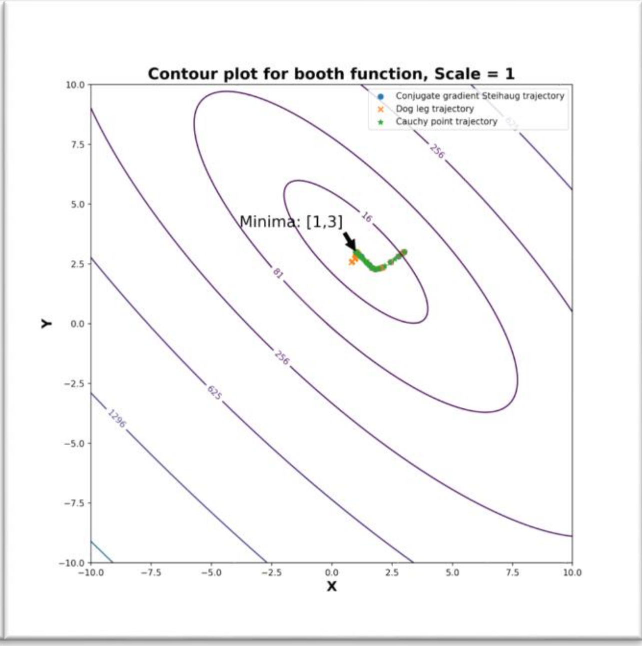
end



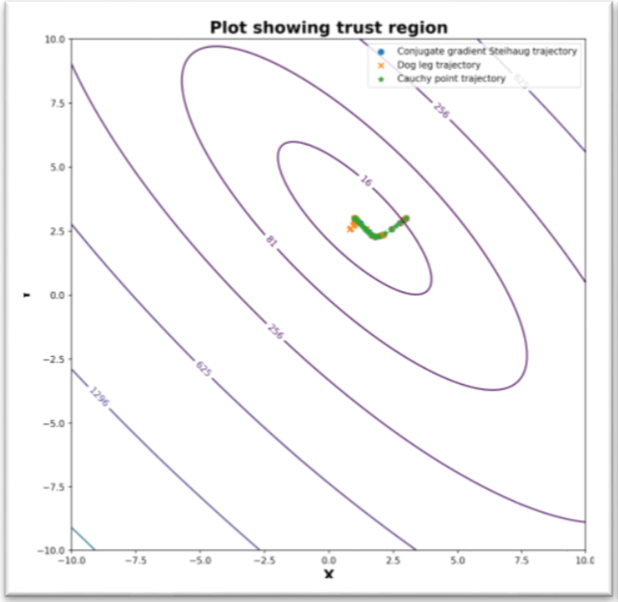
Result: Contour Plot



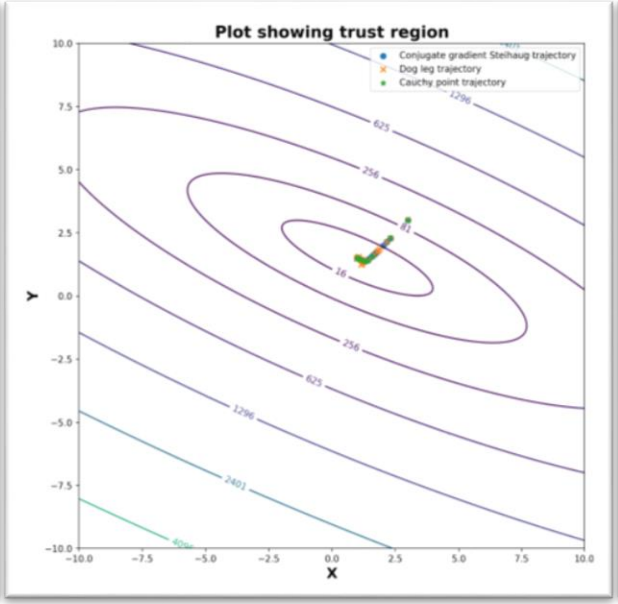




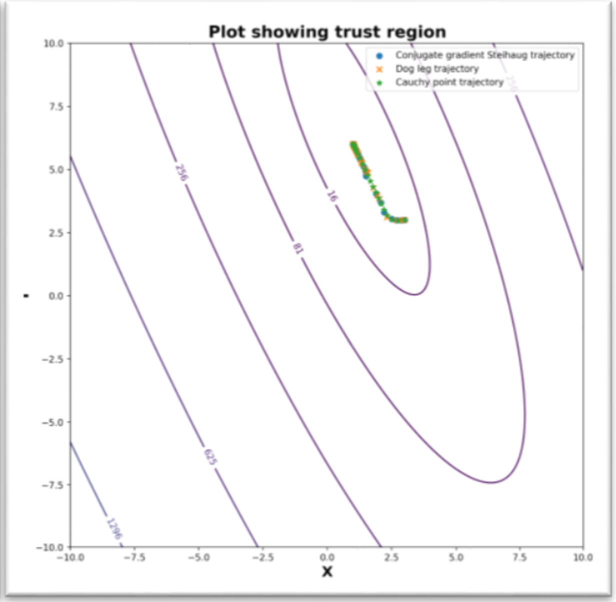
Result: Scaling (Booth function)



Scale=1.0

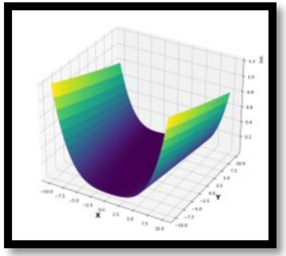


Scale=2.0



Scale=0.5

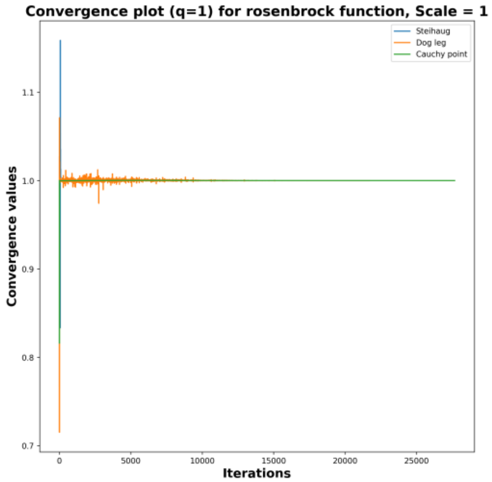
Result: Analysis (Rosenbrock)

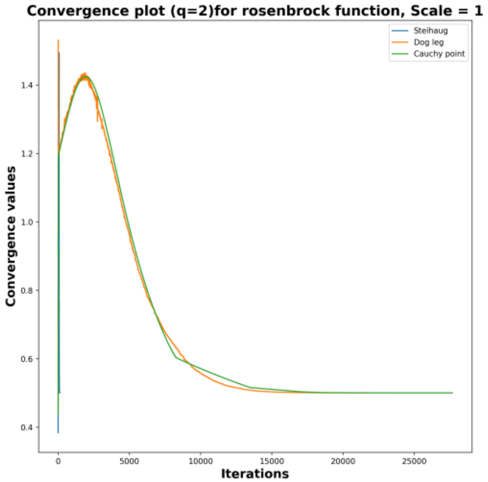


Initial parameters: $x_0 = (3, 3)$; $\Delta_{max} = 5$; $\Delta_0 = 1$

	Cauchy point	Dogleg method	CG Steihaug
Number of iterations	27678	21074	118
Execution time	2.915 sec	9.714 sec	0.035 sec
Optimal point x_k	(1.00000433, 1.00000868)	(1.0001110, 1.00022252)	(1.00007155, 1.00014345)
Scale=0.5 (iterations)	30000	30000	21
Scale=1.0 (iterations)	27678	21074	118
Scale=2.0 (iterations)	25340	22686	3301

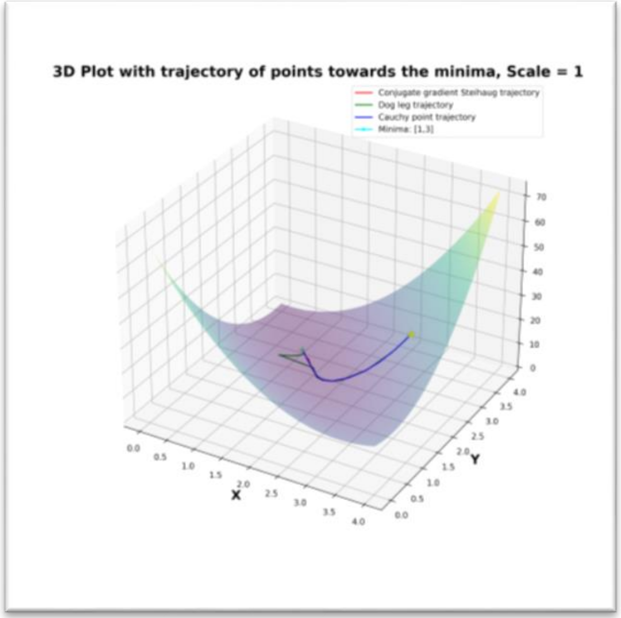
Result: Convergence (Rosenbrock)

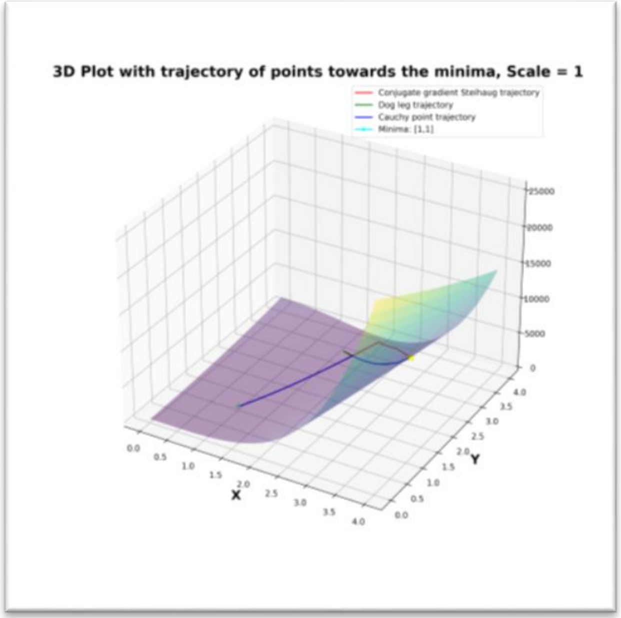


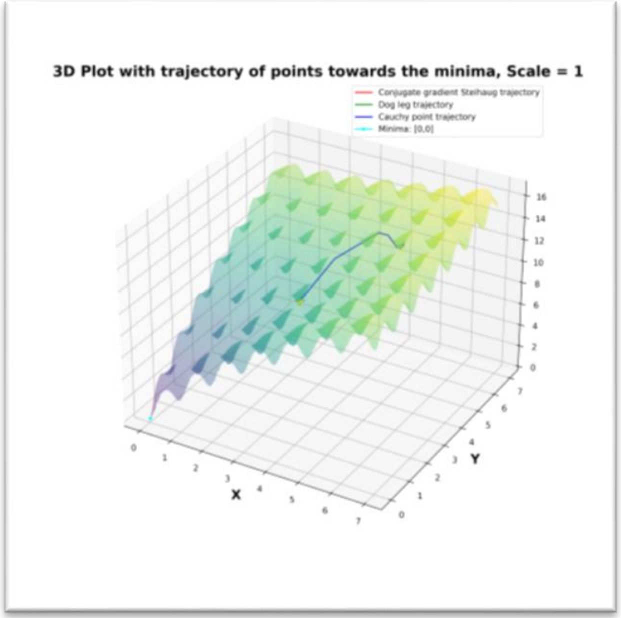


Subproblem solution methods	Rosenbrock 'q' value	Booth 'q' value	Ackley 'q' value
Cauchy point	1.000127	0.991538	0.99999
Dog Leg	1.00008	1.008220	1.00000
Steihaug	0.986129	1.011392	0.01662

Result: Trajectory 3D plot







Conclusion

- Variants to solve trust region subproblem was investigated.
- Plots describing the optimization process were made.
- Effect of scaling on the objective functions, run time of the subproblem variants were analysed.
- The convergence of subproblem solutions variants was investigated.

References

1. Numerical optimization lecture manuscript by Prof. Moritz Diehl.
2. Jorge Nocedal and Stephen J. Wright, Numerical Optimization, Springer, 2006
3. Trust region theory: https://medium.com/trust-region-methods/temp-blog-1b51189594a
4. Reference trust region code and plot functions: https://github.com/anweshpanda/Trust_Region/blob/main/Trust%20Region.ipynb
5. Steihaug method theory: https://vdocuments.mx/the-dogleg-and-steihaug-methods-science-and-technology-the-dogleg-and.html?page=22
6. Trust region theory and pseudo code: https://web.archive.org/web/20220922103636/https://optimization.mccormick.northwestern.edu/index.php/Trust-region_methods

Thank you for your attention!