Unconstrained Optimization

Presentation topics

- Methods used
- Problems analysis
- Final discussion and considerations

Objective: Explore and confirm theoretical assumptions and practical tradeoffs between methods

Methods used

Methods explanation

Methods used

Steepest Descent Method: Informed ($\nabla f(x)$ given)

Conjugate Gradient Method: Informed ($\nabla f(x)$ given)

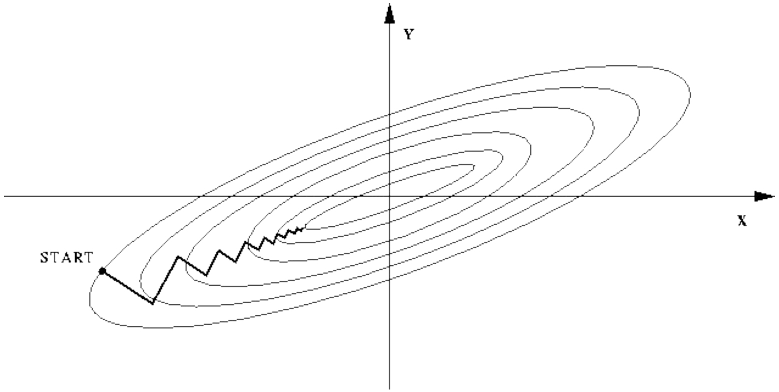
Inexact Newton Method: Uninformed ($\nabla f(x)$ and $\nabla^2 f(x)$ not given)
(with finite differences) and approximated computations

Steepest Descent Method

Takes the **negative gradient as descending direction** everytime.

Best local descent direction, not best descent direction overall.

Last direction orthogonal to the previous, thus is **conditioning dependent**.



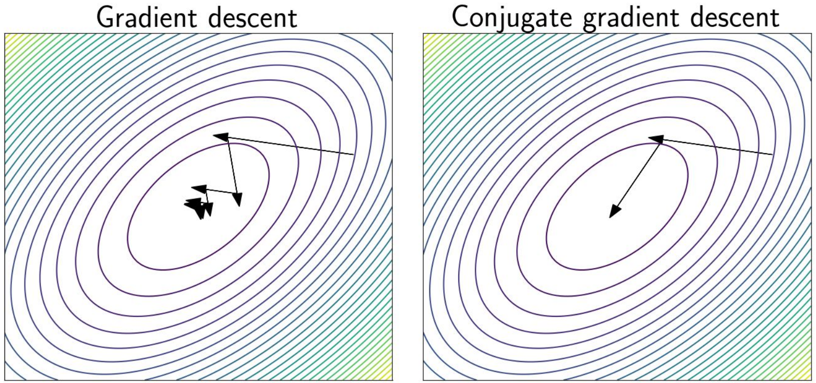
$$p_k = -\nabla f(x_k)$$

Conjugate Gradient Method

Improved SD method.

New descent direction is never orthogonal. It is conjugate to the previous.

Better convergence can be proven.



$$\beta_k = -\frac{(d_k)^T A r^{(k+1)}}{d_k{}^T A d_k}$$

$$p_{k+1} = -\nabla f_{k+1} + \beta^{k+1} p^k$$

CG method extended to a nonlinear function f:

- Steplength α computed with a line search
- Residual not anymore $r = b - Ax$, but replaced by ∇f

$$\beta^{k+1} = \frac{\nabla f_{k+1}^T \nabla f_{k+1}}{\nabla f_k^T \nabla f_k} = \frac{||\nabla f_{k+1}||^2}{||\nabla f_k||^2}$$

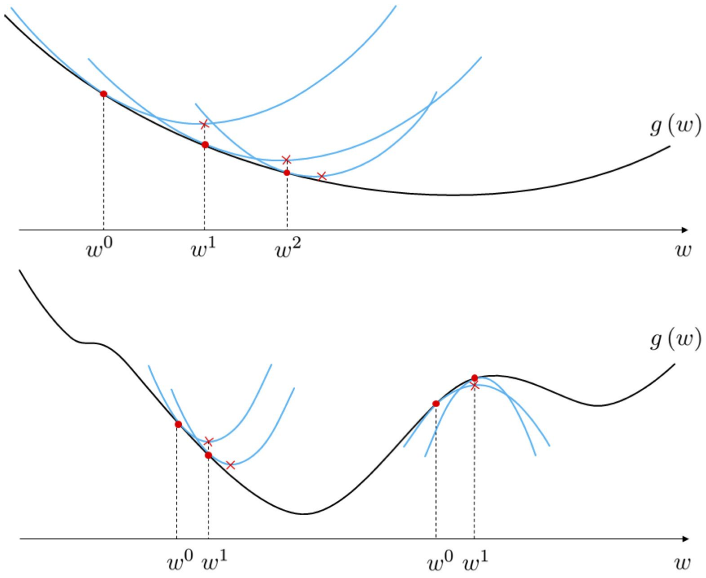
Fletcher & Reeve's version

Newton Method

Exploits the quadratic approximation of the function.

Needs both gradient and hessian of the function.

Hessian has to be SPD, orelse convergence is not guaranteed.



$$p = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k).$$

Inexact Newton Method

Exact descent direction can be hard to compute exactly (exact min of paraboloid):

$$p = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k).$$

We approximate it instead:

$$||\nabla^2 f(x_k)p + \nabla f(x_k)|| \leq \eta_k ||\nabla f(x_k)||$$

(!) Updating tolerance: Near the solution $\longrightarrow$ More precise solving

Inexact Newton Method
Forcing Terms

To compute descent direction we use an **iterative solver** (pcg).

More exact → More inner iterations (for Descent Direction)

Less exact → More outer iterations (Newton steps)

Compromise, so use a suitable **forcing term**:

$$||\nabla^2 f(x_k)p + \nabla f(x_k)|| \leq \eta_k ||\nabla f(x_k)|| \qquad \eta_k = \min(0.5, \sqrt{||\nabla f(x_k)||})$$

(superlinear convergence is not guaranteed, but generally expected)

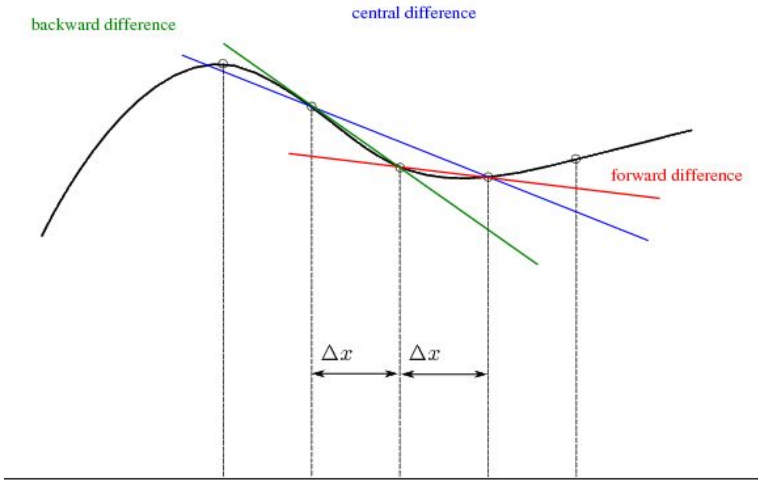
Newton Method
Finite differences

Gradient and Hessian may be hard to compute.

Approximate them with incremental ratio expressions.

We approximated both:

- gradient (central difference)
- Hessian (forward difference)

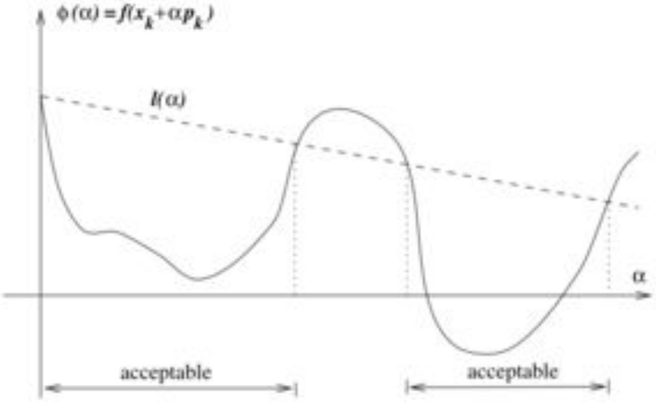


Line search and Armijo condition

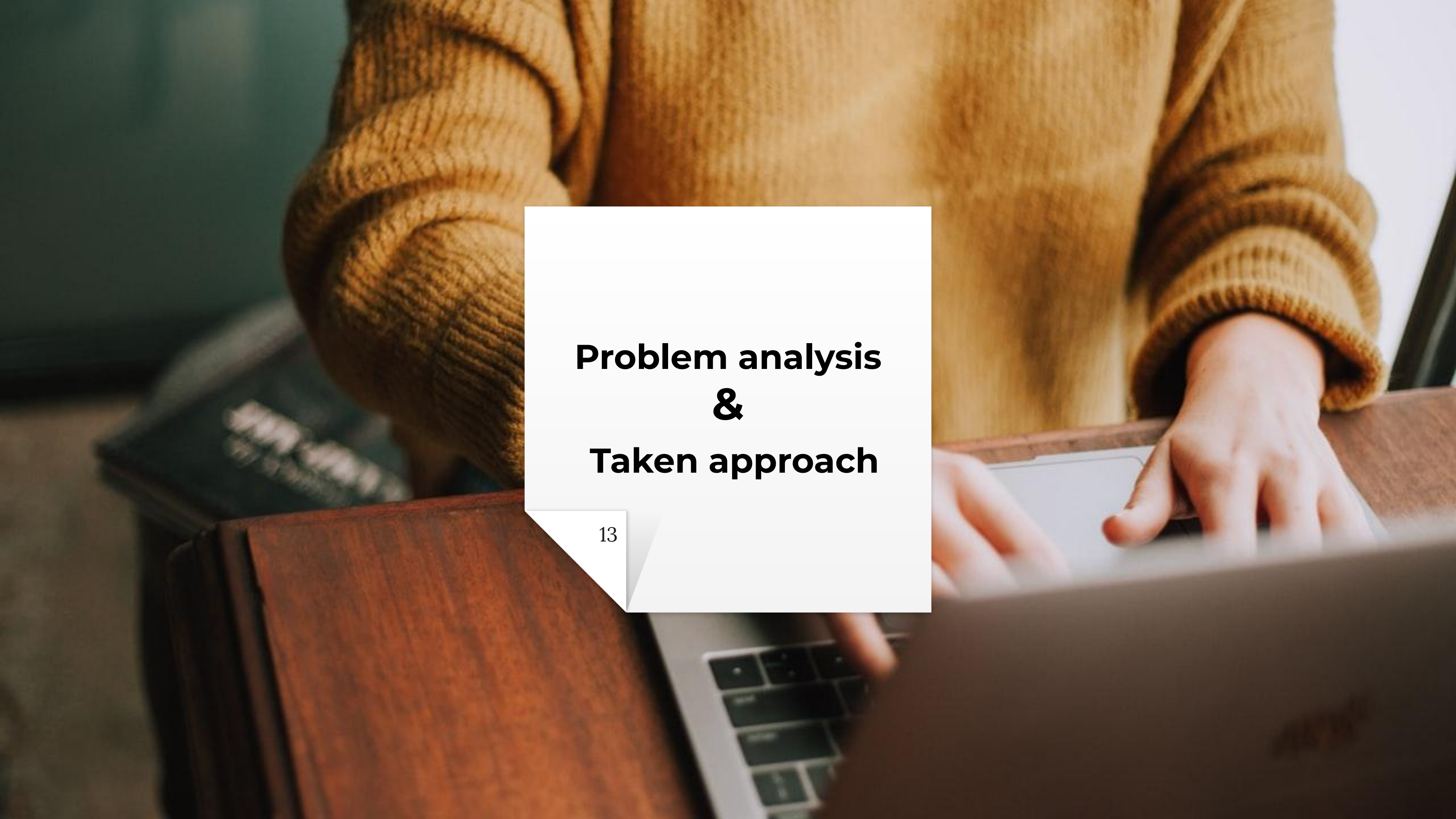
In non-linear case we cannot say apriori which is the best step length.

Only ‘descent’ is not enough.

We must choose one that guarantees ‘sufficient descent’.



$$f(\mathbf{x}_k + \alpha_k \mathbf{p}_k) \leq f(\mathbf{x}_k) + c_1 \alpha_k \mathbf{p}_k^{\mathrm{T}} \nabla f(\mathbf{x}_k)$$



Problem analysis & Taken approach

Parameters and approach

- Convergence threshold set as: $||\nabla f_k(x)|| = 10^{-6}$
- Max steps with unimprovement: 5
- Unimprovement threshold: 10^{-12}
- Armijo condition:
  - $\alpha_0=5$ ($\alpha_0=1$ for NM)
  - $\rho=0.5$
  - $C_1= 10^{-4}$
- NM, fully uninformed:
  - Gradient FD: Forward
  - Hessian FD: Central

ROSENBROCK FUNCTION

- All methods converge to the actual minimum $x^*=(1, 1)$ with starting point x_0

- Changing the start point to x_1 , SD and CG-FR perform well, while INM does not converge.

$$f(x_1, x_2) = (1 - x_1)^2 + 100\left(x_2 - x_1^2\right)^2$$

$$\overline{x_{(0)}} = (1.2, 1.2), \overline{x_{(1)}} = (-1.2, 1)$$

[P05] GENERALIZED BROYDEN TRIDIAGONAL FUNCTION

- All methods converge but convergence points are distant: problem has multiple reachable minpoints

- Changing the start point to origin they remain distant and NM does not converge ($\nabla^2 f(x)$ non positive definite)

- NM very slow (many computations for Hessian)

$$F(x) = \sum_{i=1}^{n} |(3 - 2x_i)\, x_i - x_{i-1} - x_{i+1} + 1|^p$$

$$p = 7/3, \quad \overline{x_0} = \overline{x_{n+1}} = 0$$

$$\overline{x_i} = -1, \quad i \geq 1.$$

[P13] GENERALIZATION OF THE BROWN FUNCTION 2

- All methods converge to the same point

- The convergence rate respects the theory (relatively to each other):
  - NM: 7 iterations
  - CG–FR: 10 iterations
  - SD: 15 iterations

$$F(x) = \sum_{j=1}^{k} \left[\left(x_{i-1}^2\right)^{\left(x_i^2+1\right)} + \left(x_i^2\right)^{\left(x_{i-1}^2+1\right)} \right],$$

$$i = 2j, \quad k = n/2,$$

$$\overline{x_i} = -1, \mathrm{mod}(i,2) = 1, \quad \overline{x_i} = 1, \mathrm{mod}(i,2) = 0.$$

[P14] DISCRETE BOUNDARY VALUE PROBLEM

- No method converges completely
- SD and CG-FR gets to $||\nabla f_k(x)|| = 10^{-3}$ after 50'000 iterations
- NM stopped after 5 iterations as it wasn't making progress
- Starting from the origin NM converges to the same optimal point (only better convergence criterion)

$$F(x) = \sum_{i=1}^{n} \left[2x_i - x_{i-1} - x_{i+1} + h^2 (x_i + ih + 1)^3 / 2\right]^2,$$

$$h = 1/(n+1), \quad \overline{x_0} = \overline{x_{n+1}} = 0,$$

$$\overline{x}_i = ih(1 - ih), \quad i \geq 1.$$

[P16] BANDED TRIGONOMETRIC PROBLEM

- No method converges completely
- Starting from the origin the situation slightly improves
- Trying with a smaller n (=100) they actually converge: it is only time/computing problem issue

$$F(x) = \sum_{i=1}^{n} i\left[(1 - \cos x_i) + \sin x_{i-1} - \sin x_{i+1}\right]$$

$$x_0 = x_{n+1} = 0 \qquad \bar{x}_i = 1, \quad i \geq 1$$

Final Discussion

Final Discussion

- Methods, when applicable, do respect theoretical convergence rate (at least in respect to each other)
- Fully uninformed NM has too many computations every iteration, thus takes a lot of time to finish
- NM is unreliable as it's hessian matrix can become non-PD and this can change based on starting point
    - ***Improvement***: Changing Hessian to become positive definite at each iteration can solve the non-reliability problem

Team Presentation

Florentin-Cristian Udrea
S319029

Sara Rosato
S317547

Thanks for the attention!