Differential programming

Sergey Barseghyan barseghyan@phystech.edu

16.08.2022

Outline

1. Non-linear optimization and SABR model
   - Levenberg-Marquad
   - SABR model
2. Diffential programming for detectors

Levenberg-Marquad algorithm - is an optimization algorithm which is basically combination of two objective minimization algorithms: **Gauss-Newton** and **Gradient Descent** It is very very well suited for the optimization problemsswhere model is **non-linear** in its parameters

Non-linear least squares

$$\begin{aligned}
\hat{y}(t;\boldsymbol{p}) &\quad \text{- fitted model} \\
t \in \mathbb{R} &\quad \text{- independent variable} \\
\{t_i, y_i\}_{i=1}^{m} &\quad \text{- observed data points} \\
W_{ij} &\quad \text{- inverse covariance matrix} \\
\boldsymbol{p} \in \mathbb{R}^n &\quad \text{model parameters}
\end{aligned}$$

Optimized objective is nothing else but chi-squared statistic

$$\chi^2(\boldsymbol{p}) = \sum_{i=1}^{m} \left[\frac{y(t_i) - \hat{y}(t_i;\boldsymbol{p})}{\sigma_{y_i}} \right]^2$$

$$= (\boldsymbol{y} - \hat{\boldsymbol{y}}(\boldsymbol{p}))^\top \boldsymbol{W} (\boldsymbol{y} - \hat{\boldsymbol{y}}(\boldsymbol{p}))$$

Gradient Descent and Jacobian update

Gradient Descent step update

$$\frac{\partial}{\partial \boldsymbol{p}}\chi^2 = -2(\boldsymbol{y} - \hat{\boldsymbol{y}})^\top \boldsymbol{W}\boldsymbol{J}$$

$$\boldsymbol{h}_{\mathrm{gd}} = \alpha \boldsymbol{J}^\top \boldsymbol{W}(\boldsymbol{y} - \hat{\boldsymbol{y}})$$

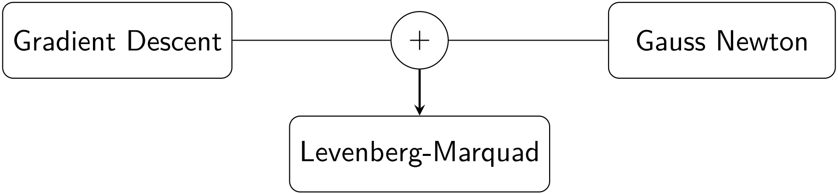
Gauss-Newton step update

$$\hat{\boldsymbol{y}}(\boldsymbol{p} + \boldsymbol{h}) \approx \hat{\boldsymbol{y}}(\boldsymbol{p}) + \left[\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{p}}\right] \boldsymbol{h} = \hat{\boldsymbol{y}} + \boldsymbol{J}\boldsymbol{h}$$

$$\frac{\partial}{\partial \boldsymbol{h}}\chi^2(\boldsymbol{p} + \boldsymbol{h}) \approx -2(\boldsymbol{y} - \hat{\boldsymbol{y}})^\top \boldsymbol{W}\boldsymbol{J} + 2\boldsymbol{h}^\top \boldsymbol{J}^\top \boldsymbol{W}\boldsymbol{J}$$

$$\left[\boldsymbol{J}^\top \boldsymbol{W}\boldsymbol{J}\right] \boldsymbol{h}_{\mathrm{gn}} = \boldsymbol{J}^\top \boldsymbol{W}(\boldsymbol{y} - \hat{\boldsymbol{y}})$$

Levenberg-Marquad update



$$\left[\boldsymbol{J}^\top \boldsymbol{W}\boldsymbol{J} + \lambda\, \mathrm{diag}\left(\boldsymbol{J}^\top \boldsymbol{W}\boldsymbol{J}\right)\right] \boldsymbol{h}_{\mathrm{lm}} = \boldsymbol{J}^\top \boldsymbol{W}(\boldsymbol{y} - \hat{\boldsymbol{y}})$$

λ is dumping parameter

Motivation for the model

Model Definition

Exact solution