Quasi-Newton methods (BFGS)

Quasi-Newton methods, like Broyden–Fletcher–Goldfarb–Shanno (BFGS), are a family of optimization algorithms designed to address limitations of standard gradient descent in finding minima of functions.

Quasi-Newton Methods (QN):

- Aim to overcome the limitations of gradient descent by incorporating curvature information.
- They don't calculate the Hessian (second derivative) directly, which can be expensive, but approximate it using past gradients and function evaluations.
- This allows them to take larger steps towards the minimum compared to gradient descent, potentially leading to faster convergence.

Advantages:

- **Faster Convergence:** Compared to gradient descent, BFGS can converge to the minimum point much faster, especially for functions with non-quadratic curvatures.
- **Efficient Hessian Approximation:** It avoids the expensive calculation of the full Hessian by using a low-rank update scheme for the BFGS matrix.
- **Widely Used:** BFGS is a popular choice for various optimization problems due to its efficiency and good performance.

Disadvantages to:

- **Memory Requirements:** Storing the BFGS matrix can require more memory compared to gradient descent, especially for high-dimensional problems.
- **Convergence Guarantees:** Unlike Newton's method, BFGS doesn't have guaranteed convergence for all functions. However, it often performs well in practice.