Conjugate Gradient Method for Optimization

Ankesh Kumar

1 Introduction to Conjugate Direction Methods

Conjugate direction methods are a class of optimization algorithms designed to minimize quadratic functions efficiently. These methods generate a set of search directions that are conjugate with respect to the Hessian matrix of the objective function.

2 Conjugate Gradient Algorithm for General Functions

The conjugate gradient method can be extended to non-quadratic functions. The general algorithm is as follows:

1. Initialize x_0, compute $g_0 = \nabla f(x_0)$, set $d_0 = -g_0$
2. For $k = 0, 1, 2, \ldots$ until convergence:
   (a) Perform line search: $\alpha_k = \arg\min_\alpha f(x_k + \alpha d_k)$
   (b) Update: $x_{k+1} = x_k + \alpha_k d_k$
   (c) Compute new gradient: $g_{k+1} = \nabla f(x_{k+1})$
   (d) Compute β_k using one of the following formulas:
      - Fletcher-Reeves (FR): $\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$
      - Hestenes-Stiefel (HS): $\beta_k^{HS} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T (g_{k+1} - g_k)}$
      - Polak-Ribière (PR): $\beta_k^{PR} = \max\left\{0, \frac{g_{k+1}^T (g_{k+1} - g_k)}{g_k^T g_k}\right\}$

   (e) Update search direction: $d_{k+1} = -g_{k+1} + \beta_k d_k$

3 Results and Discussion

We applied the conjugate gradient method to the Rosenbrock function:

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2$$

Starting from the initial point $(-2, 2)$, we obtained the following results:

- Fletcher-Reeves update:
  - Optimal solution: $(-1.9062, 2.0239)$
  - Optimal function value: 267.5601
- Hestenes-Stiefel update:
  - Optimal solution: $(0.99999, 0.99998)$
  - Optimal function value: 9.7528×10^{-11}
- Polak-Ribière update:
  - Optimal solution: $(1, 1)$
  - Optimal function value: 5.0124×10^{-13}

The results show that the Polak-Ribière formula performed best, converging to the global minimum $(1, 1)$ with high accuracy. The Hestenes-Stiefel update also performed well, reaching a point very close to the global minimum. The Fletcher-Reeves update, however, did not converge to the global minimum in this case, possibly due to the challenging nature of the Rosenbrock function's landscape.

These results highlight the importance of choosing an appropriate β_k formula for the conjugate gradient method, especially when dealing with non-quadratic functions like Rosenbrock. The Polak-Ribière formula, with its built-in restart capability (when β_k is set to 0), seems to be particularly effective for this problem.