NBA Win Percentage Prediction: Comparison of Optimization Methods for Ridge Regression

OBJECTIVE: Minimizing Ridge Regression Loss

The primary objective is to predict NBA win percentages by determining the optimal coefficients β for a Ridge Regression model. This involves minimizing the following loss function.

$$L(\beta) = \|y - X\beta\|^2 + \lambda\|\beta\|^2$$

Data Set

The research utilizes extensive NBA team statistics from the NBA Team Stats dataset (2000–2018) from Kaggle. This comprehensive dataset includes regular season and playoff statistics for NBA teams, featuring wins, losses, points, assists per game, turnovers per game, total rebounds per game, field goal percentage, and three-point percentage.

Methods

1. Ridge Regression Loss Function

$$L(\beta) = \|y - X\beta\|^2 + \lambda\|\beta\|^2$$

2. Steepest Descent Update

$$\beta_{k+1} = \beta_k - \alpha\nabla L(\beta_k)$$

3. Conjugate Gradient Update

$$d_{k+1} = -\nabla L(\beta_{k+1}) + \gamma_k d_k$$

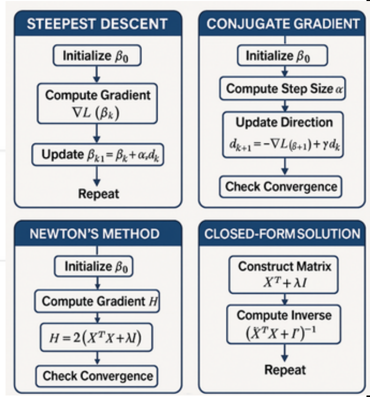
$$\gamma_k = \frac{\|\nabla L(\beta_{k+1})\|^2}{\|\nabla L(\beta_k)\|^2}$$

4. Newton's Method Update

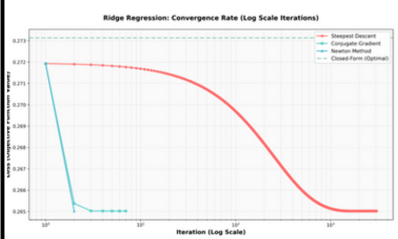
$$\beta_{k+1} = \beta_k - H^{-1}\nabla L(\beta_k)$$

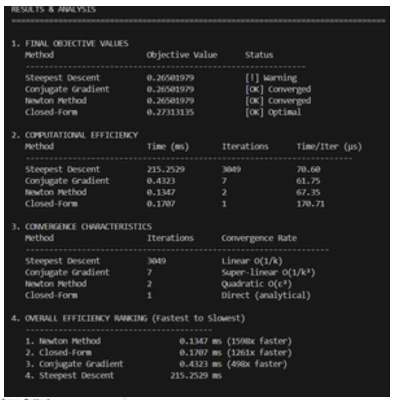
with Hessian

$$H = 2(X^TX + \lambda I)$$



Results





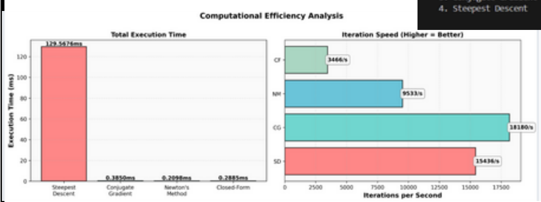
```
RESULTS & ANALYSIS
================================================================================

1. FINAL OBJECTIVE VALUES
   Method                    Objective Value     Status
   ------------------------------------------------------------
   Steepest Descent          0.26501979          [!] Warning
   Conjugate Gradient        0.26501979          [OK] Converged
   Newton Method             0.26501979          [OK] Converged
   Closed-Form               0.27313135          [OK] Optimal

2. COMPUTATIONAL EFFICIENCY
   Method                    Time (ms)       Iterations      Time/Iter (μs)
   ------------------------------------------------------------------------
   Steepest Descent          215.2529        3049            70.60
   Conjugate Gradient        0.4323          7               61.75
   Newton Method             0.1347          2               67.35
   Closed-Form               0.1707          1               170.71

3. CONVERGENCE CHARACTERISTICS
   Method                    Iterations      Convergence Rate
   ------------------------------------------------------------------
   Steepest Descent          3049            Linear O(1/k)
   Conjugate Gradient        7               Super-linear O(1/k²)
   Newton Method             2               Quadratic O(ε²)
   Closed-Form               1               Direct (analytical)

4. OVERALL EFFICIENCY RANKING (Fastest to Slowest)
   ----------------------------------------
   1. Newton Method             0.1347 ms (1598x faster)
   2. Closed-Form               0.1707 ms (1261x faster)
   3. Conjugate Gradient        0.4323 ms (498x faster)
   4. Steepest Descent        215.2529 ms
```



Conclusion

- All methods converge to the same optimal Ridge Regression objective value (0.26), confirming correctness across approaches.
- Steepest Descent converges slowly (3049 iterations) and does not fully reach the optimum within the iteration limit.
- Newton's Method achieves the fastest convergence (2 iterations) due to quadratic convergence, but requires higher per-iteration computation.
- The Closed-Form solution provides an exact answer in a single step, serving as a useful benchmark but scaling poorly to high dimensions.
- Conjugate Gradient offers the best overall trade-off, reaching the optimal solution in few iterations with low computational cost, making it the most practical method for large datasets.