Exploring Optimization Methods in GLM Across Different Frameworks

Generalized Linear Models (GLMs) are widely used in statistical modeling, and different frameworks implement distinct optimization techniques to estimate parameters efficiently. Below is a comparative summary of the optimization techniques across various frameworks.

Base R (stats package) Reweighted Least Squares approach that updates weights iteratively to estimate GLM parameters. GLMs apply without regularization.
Squares (IRLS) Squares (IRLS) Updates weights iteratively to estimate GLMs apply without regularization. Big Data R (High-Performance Computing R) Parallelized IRLS Similar to IRLS but optimized for distributed computing. Stochastic Gradient Descent (SGD) Spark R (MLlib in Apache Spark) Spark Optimization Squares (IRLS) Updates weights iteratively to estimate GLMs apply without regularization. Similar to IRLS but optimized for distributed computing. Similar to IRLS but optimized for distributed computing environments. Iteratively updates model parameters using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Implements different Similar to IRLS but optimized for large-scale GLM problems in high-performance computing environments. Ideal for handling large datasets that do not fit in memory. Preferred for big data applications requiring distributed computation.
Big Data R (High-Performance Computing R)
Big Data R (High- Performance Computing R) Parallelized IRLS Computing R) Similar to IRLS but optimized for distributed computing. Iteratively updates model parameters using gradients to minimize the error. Spark R (MLlib in Apache Spark) Spark Optimization Spark Optimization Gradient Descent GLM parameters. Similar to IRLS but optimized for distributed computing. Iteratively updates model parameters using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Figuralization. Similar to IRLS but optimized for distributed acmputing environments. Ideal for handling large datasets that do not fit in memory. Preferred for big data applications requiring distributed computation.
Big Data R (High-Performance Computing R)Parallelized IRLSSimilar to IRLS but optimized for distributed computing.Best suited for large-scale GLM problems in high-performance computing environments.Dask MLStochastic Gradient Descent (SGD)Iteratively updates model parameters using gradients to minimize the error.Ideal for handling large datasets that do not fit in memory.Spark R (MLlib in Apache Spark)LBFGS (Limited-memory BFGS)An optimization algorithm for convex functions, useful for high-dimensional data.Preferred for big data applications requiring distributed computation.Spark OptimizationGradient DescentImplements differentEffective in large-
Performance Computing R) Dask ML Stochastic Gradient Descent (SGD) Spark R (MLlib in Apache Spark) Spark Optimization Spark Optimization Problems in high-performance computing. Iteratively updates model parameters using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Gradient Descent Spark Optimization Gradient Descent Spark Optimization Gradient Descent Spark Optimization Gradient Descent Optimized for distributed computation Ideal for handling large datasets that do not fit in memory. Preferred for big data applications requiring distributed computation.
Computing R) Dask ML Stochastic Gradient Descent (SGD) Spark R (MLlib in Apache Spark) Spark Optimization Spark Optimization Gradient Descent Computing. Iteratively updates model parameters using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Ideal for handling large datasets that do not fit in memory. Preferred for big data applications requiring distributed computation.
Dask ML Stochastic Gradient Descent (SGD) Spark R (MLlib in Apache Spark) Spark Optimization Computing. Iteratively updates model parameters using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Computing. Computing. Ideal for handling large datasets that do not fit in memory. Preferred for big data applications requiring distributed computation.
Dask ML Stochastic Gradient Descent (SGD) Spark R (MLlib in Apache Spark) LBFGS (Limited- memory BFGS) Brank Optimization Gradient Descent Stochastic Gradient Descent (SGD) LBFGS (Limited- memory BFGS) An optimization algorithm for convex functions, useful for high-dimensional data. Effective in large-
Stochastic Gradient Descent (SGD) Iteratively updates model parameters using gradients to minimize the error. Ideal for handling large datasets that do not fit in memory.
Stochastic Gradient Descent (SGD) Spark R (MLlib in Apache Spark) Spark Optimization Stochastic Gradient Descent (SGD) Stochastic Gradient Descent (SGD) Model parameters using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Stochastic Gradient Descent model parameters using gradients to mot fit in memory. Preferred for big data applications requiring distributed computation.
Spark R (MLlib in Apache Spark) Spark Optimization Consider the content of the
Spark R (MLlib in Apache Spark) Spark R (MLlib in Apache Spark) Spark Optimization Gradient Descent (SGD) Using gradients to minimize the error. An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Using gradients to mot fit in memory. Preferred for big data applications requiring distributed computation.
Spark R (MLlib in Apache Spark) LBFGS (Limited-memory BFGS) An optimization algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent Implements different Effective in large-
Apache Spark) memory BFGS) algorithm for convex functions, useful for high-dimensional data. Spark Optimization Gradient Descent applications requiring distributed computation. Effective in large-
functions, useful for high-dimensional data. Spark Optimization Gradient Descent functions, useful for high-dimensional computation. Implements different Effective in large-
high-dimensional computation. Spark Optimization Gradient Descent Implements different Effective in large-
Spark Optimization Gradient Descent Implements different Effective in large-
Spark Optimization Gradient Descent Implements different Effective in large-
(MLIID) variants lorms of gradient scale machine
descent for learning workflows
optimizing GLM across multiple
parameters. nodes.
Scikit-learn Coordinate Descent A regularization- Best suited for high-
& Elastic Net based method that dimensional sparse
updates coefficients datasets, such as text
iteratively to improve classification or
prediction accuracy. genomic data.

In brief IRLS remains a classic choice for traditional GLMs, but when handling large datasets, parallelized or distributed methods (e.g., Spark's LBFGS) become essential. Gradient-based methods (SGD, LBFGS) are powerful when dealing with vast amounts of data that require scalable solutions. Regularization methods (Elastic Net, Coordinate Descent) shine in high-dimensional data where feature selection is crucial.