GENO – GENeric Optimization for Classical Machine Learning

Sören Laue    Matthias Mitterreiter    Joachim Giesen

Friedrich-Schiller-University Jena, Germany

NeurIPS 2019
Machine Learning

About 600-2000 papers published at NeurIPS and ICML per year.

~25% design / implement a new optimization algorithm
Look at machine learning through the lens of optimization.

- least squares regression
- logistic regression
- k-means
- network analytics
- (hyperbolic) embeddings
- …
Work flow

1. Model problem
2. Implement solver
3. Evaluate solution
General Optimization in Machine Learning

Ideal world:

▶ One tool / algorithm for everything
▶ Easy to use
▶ As fast as hand-tuned, specialized solvers
Our approach
matrix A
vector b
minimize norm(A*x - b, 2)
subject to
x >= 0

matrix X sparse
vector y
scalar c
minimize
1/2 * w'*w + c * sum(xi)
subject to
y.*(X*w+vector(b)) >= 1-xi
xi >= 0

non-negative
least squares
solver

SVM solver
Results
# Kernelized Dual SVM

<table>
<thead>
<tr>
<th>data set</th>
<th>LIBSVM</th>
<th>GENO</th>
<th>CVX / SeDuMi</th>
<th>CVX / Gurobi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sec.</td>
<td>Sec.</td>
<td>Sec.</td>
<td>Sec.</td>
</tr>
<tr>
<td>a1a</td>
<td>0.28</td>
<td><strong>0.28</strong></td>
<td>376.6</td>
<td>57.8</td>
</tr>
<tr>
<td>a7a</td>
<td>39.5</td>
<td><strong>29.8</strong></td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

LIBSVM – Chang and Lin 2001
## Logistic Regression

<table>
<thead>
<tr>
<th>data set</th>
<th>LIBLINEAR</th>
<th>GENO</th>
<th>CVX / SeDuMi</th>
<th>CVX / Gurobi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sec.</td>
<td>Sec.</td>
<td>Sec.</td>
<td>Sec.</td>
</tr>
<tr>
<td>a1a</td>
<td>0.01</td>
<td>0.01</td>
<td>254.3</td>
<td>n/a</td>
</tr>
<tr>
<td>rcv1_test</td>
<td>12.8</td>
<td>5.5</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

LIBLINEAR – Lin et al. (JMLR 2008)
Elastic Net Regression

<table>
<thead>
<tr>
<th>data set</th>
<th>glmnet</th>
<th>GENO</th>
<th>CVX/Gurobi</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1000, 1000)</td>
<td>0.10</td>
<td><strong>0.11</strong></td>
<td>21.1</td>
</tr>
<tr>
<td>(5000, 10000)</td>
<td>4.75</td>
<td><strong>4.11</strong></td>
<td>n/a</td>
</tr>
</tbody>
</table>

glmnet – Friedman, Hastie, Tibshirani (JStatSoft 2010)
Sparse PCA (non-linear version)

<table>
<thead>
<tr>
<th>data set</th>
<th>GPower</th>
<th>GENO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fValue</td>
<td>Sec.</td>
</tr>
<tr>
<td>colon-cancer</td>
<td>-72.3</td>
<td>0.0109</td>
</tr>
<tr>
<td>gisette</td>
<td>-34.5</td>
<td>1.98</td>
</tr>
</tbody>
</table>

GPower – Journée et al. (JMLR 2010)
Detailed Results
SAGA (NIPS 2014), CSDCA (NIPS 2015), Point-SAGA (NIPS 2016), LIBLINEAR (JMLR 2008)
Symmetric Non-negative Matrix Factorization

SymANLS, SymHALS (NeurIPS 2018)
Compressed Sensing

IRLS (ICML 2019)
Results:

- One tool / algorithm for everything a lot
- Easy to use
- No tuning needed
- As fast as hand-tuned, specialized, well-established solvers
- Outperforms state-of-the-art, recently published solvers by a large margin

geno-project.org

Sören Laue, Matthias Mitterreiter, Joachim Giesen.