The University of Texas
at Austin

EE381V Large Scale Optimization

Problem Set 0

Edited by LaTeX

Department of Computer Science

STUDENT

Jimmy Lin
xl5224

COURSE COORDINATOR

Sujay Sanghavi

UNIQUE NUMBER

17350

RELEASE DATE

September 5, 2014

DUE DATE

September 11, 2014

TIME SPENT

10 hours

September 7, 2014

Table of Contents

Chapter 1

Matlab and Computational Assignment

1.1 Sparse Recovery

1.1.1 Algorithm 1: Least Square

Small-scale dataset: Succeed

Total CPU time (secs) = 0.18
CPU time per iteration = 0.02
Regression Error: $+1.48902e^{-10}$
Testing Error: 23.058394

Medium-scale dataset: Succeed

Total CPU time (secs) = 43.95
CPU time per iteration = 5.49
Regression Error: $+5.92177e^{-10}$
Testing Error: 19.862394

Large-scale dataset: Failed

The scale of this least-square task is too large.

1.1.2 Algorithm 2: optimization with LASSO

Small-scale dataset: Succeed

Total CPU time (secs) = 0.38
CPU time per iteration = 0.02
Regression Error: +5.02333
Testing Error: 0.144338
Support(non-zeros entries of β): 43

Medium-scale dataset: Succeed

Total CPU time (secs) = 126.66
CPU time per iteration = 4.87
Regression Error: +5.11426
Testing Error: 0.078289
Support (non-zeros entries of β): 342

Large-scale dataset: Failed

The scale of this least-square task is too large.

1.2 Orthogonal Matching Pursuit

Chapter 2

Linear Algebra Review

Appendix A

Codes Printout

A.1 Sparse Recovery

A.1.1 Algorithm 1: Least Square

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Scripts invoking cvx least-square routines to
%%% solve problems using our three datasets.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% solve least-square problem for first dataset
cvx_begin
    variable b1(size(X1,2))
    minimize( norm( X1*b1-y1 ) )
cvx_end
%% apply learned model on testing data
pred1 = X1test * b1;

RegressionError1 = norm( X1*b1-y1 )
TestingError1 = norm(pred1 - y1test)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% solve least-square problem for second dataset
cvx_begin
    variable b2(size(X2,2))
    minimize( norm( X2*b2-y2 ) )
cvx_end
%% apply learned model on testing data
pred2 = X2test * b2;

RegressError2 = norm( X2*b2-y2 )
TestError2 = norm(pred2 - y2test)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% solve least-square problem for second dataset
cvx_begin
    variable b3(size(X3,2))
    minimize( norm( X3*b3-y3 ) )
cvx_end
%% apply learned model on testing data
pred3 = X3test * b3;

RegressionError3 = norm( X3*b3-y3 )
TestingError3 = norm(pred3 - y3test)
```

A.1.2 Algorithm 2: optimization with LASSO

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% Scripts invoking cvx least-square routines to
%%% solve LASSO problems using our three datasets.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

format short e
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% solve least-square problem for first dataset
cvx_begin
    variable b1(size(X1,2))
    minimize( norm( X1*b1-y1 ) + norm(b1,1) )
cvx_end
%% apply learned model on testing data
pred1 = X1test * b1;

RegressionError1 = norm( X1*b1-y1 )
TestingError1 = norm(pred1 - y1test)
Support1 = sum(((b1 < 10e-5) + (b1 > -10e-5)) < 2)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% solve least-square problem for second dataset
cvx_begin
    variable b2(size(X2,2))
    minimize( norm( X2*b2-y2 ) + norm(b2, 1))
cvx_end
%% apply learned model on testing data
pred2 = X2test * b2;

RegressionError2 = norm( X2*b2-y2 )
TestingError2 = norm(pred2 - y2test)
Support2 = sum(((b2 < 10e-5) + (b2 > -10e-5)) < 2)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% solve least-square problem for second dataset
cvx_begin
    variable b3(size(X3,2))
    minimize( norm( X3*b3-y3 ) + norm(b3, 1) )
cvx_end
%% apply learned model on testing data
pred3 = X3test * b3;

RegressionError3 = norm( X3*b3-y3 )
TestingError3 = norm(pred3 - y3test)
Support3 = sum(((b3 < 10e-5) + (b3 > -10e-5)) < 2)
```

A.2 Orthogonal Matching Pursuit

Appendix B

CVX Experiment Results: Sparse Recovery

B.1 Algorithm 1: Least Square

B.1.1 Small-scale dataset

```
    Calling SDPT3 4.0: 551 variables, 50 equality constraints
------------------------------------------------------------

 num. of constraints = 50
 dim. of socp   var  = 51,   num. of socp blk  =  1
 dim. of free   var  = 500 *** convert ublk to lblk
*****************************************************************
   SDPT3: Infeasible path-following algorithms
*****************************************************************
 version  predcorr  gam  expon  scale_data
    NT      1      0.000   1        0
it pstep dstep pinfeas dinfeas  gap      prim-obj      dual-obj    cputime
-----------------------------------------------------------------
 0|0.000|0.000|9.5e-01|1.7e+04|1.1e+09| 2.608637e+01  0.000000e+00| 0:0:00| chol  1  1
 1|1.000|0.989|1.4e-10|1.9e+02|2.4e+06| 1.929302e+01  3.828915e-04| 0:0:00| chol  1  1
 2|0.995|0.990|2.0e-10|2.1e+00|5.2e+03| 2.098412e-01  3.092593e-05| 0:0:00| chol  1  1
 3|0.991|1.000|1.8e-11|7.5e-02|3.5e+01| 2.098816e-03  3.804833e-05| 0:0:00| chol  1  1
 4|1.000|1.000|5.8e-11|7.5e-03|5.8e-01| 2.682428e-05  2.478894e-05| 0:0:00| chol  1  1
 5|1.000|1.000|2.3e-11|7.5e-04|9.8e-03| 4.595280e-07  1.514185e-05| 0:0:00| chol  1  1
 6|1.000|1.000|9.1e-12|7.4e-05|2.1e-04| 7.745935e-09  8.345387e-06| 0:0:00| chol  1  1
 7|1.000|0.907|5.7e-12|3.2e-06|1.9e-06| 2.420707e-08  2.968077e-06| 0:0:00| chol  1  1
 8|0.990|0.988|1.1e-11|3.5e-08|1.3e-08| 2.821459e-10  3.459447e-08| 0:0:00| chol  1  1
 9|0.473|0.945|6.0e-12|1.8e-09|2.3e-10| 1.489021e-10  1.919896e-09| 0:0:00|
  stop: max(relative gap, infeasibilities) < 1.49e-08
-----------------------------------------------------------------
 number of iterations   =  9
 primal objective value =  1.48902080e-10
 dual   objective value =  1.91989619e-09
 gap := trace(XZ)       = 2.29e-10
 relative gap           = 2.29e-10
 actual relative gap    = -1.77e-09
 rel. primal infeas     = 6.03e-12
 rel. dual   infeas     = 1.78e-09
 norm(X), norm(y), norm(Z) = 7.8e-01, 1.1e-10, 1.0e+00
 norm(A), norm(b), norm(C) = 2.2e+02, 1.9e+01, 2.0e+00
```

```
 Total CPU time (secs)  = 0.18
 CPU time per iteration = 0.02
 termination code       =  0
 DIMACS: 1.6e-11  0.0e+00  1.8e-09  0.0e+00  -1.8e-09  2.3e-10
------------------------------------------------------------------
------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +1.48902e-10

Testing Error: 23.058394
```

B.1.2 Medium-scale dataset

```
    Calling SDPT3 4.0: 5501 variables, 500 equality constraints
------------------------------------------------------------


 num. of constraints = 500
 dim. of socp   var  = 501,   num. of socp blk  =  1
 dim. of free   var  = 5000 *** convert ublk to lblk
*****************************************************************
   SDPT3: Infeasible path-following algorithms
*****************************************************************
 version  predcorr  gam  expon  scale_data
    NT      1      0.000   1        0
it pstep dstep pinfeas dinfeas  gap      prim-obj      dual-obj    cputime
-----------------------------------------------------------------
 0|0.000|0.000|9.8e-01|5.2e+03|1.0e+08| 1.077392e+02  0.000000e+00| 0:0:05| chol  1  1
 1|1.000|0.989|2.0e-11|5.7e+01|2.3e+05| 9.823617e+01  6.299850e-05| 0:0:10| chol  1  1
 2|1.000|0.992|1.6e-11|6.1e-01|4.9e+02| 3.633210e+00  3.969663e-05| 0:0:15| chol  1  1
 3|0.989|1.000|7.9e-11|1.5e-02|2.3e+00| 3.834024e-02  2.056260e-05| 0:0:20| chol  1  1
 4|0.989|1.000|8.3e-11|1.5e-03|3.8e-02| 4.202595e-04  1.294517e-05| 0:0:25| chol  1  1
 5|0.989|1.000|3.4e-11|9.2e-05|4.4e-04| 4.617135e-06  7.465648e-06| 0:0:30| chol  1  1
 6|0.982|0.858|3.1e-11|3.5e-06|1.6e-06| 9.312820e-08  3.340703e-06| 0:0:34| chol  1  1
 7|0.988|0.988|1.2e-10|4.0e-08|4.5e-09| 1.122080e-09  4.025133e-08| 0:0:39| chol  1  1
 8|0.472|0.945|6.3e-11|2.2e-09|6.1e-10| 5.921772e-10  2.233900e-09| 0:0:44|
  stop: max(relative gap, infeasibilities) < 1.49e-08
-----------------------------------------------------------------
 number of iterations   =  8
 primal objective value =  5.92177202e-10
 dual   objective value =  2.23390030e-09
 gap := trace(XZ)       = 6.08e-10
 relative gap           = 6.08e-10
 actual relative gap    = -1.64e-09
 rel. primal infeas     = 6.30e-11
 rel. dual   infeas     = 2.23e-09
 norm(X), norm(y), norm(Z) = 7.2e-01, 4.6e-11, 1.0e+00
 norm(A), norm(b), norm(C) = 2.2e+03, 5.2e+01, 2.0e+00
 Total CPU time (secs)  = 43.95
 CPU time per iteration = 5.49
 termination code       =  0
 DIMACS: 3.4e-10  0.0e+00  2.2e-09  0.0e+00  -1.6e-09  6.1e-10
------------------------------------------------------------------
------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +5.92177e-10
```

```
Testing Error: 19.862394
```

B.2 Algorithm 2: Optimization with LASSO

B.2.1 Small-scale dataset

```
Calling SDPT3 4.0: 1051 variables, 50 equality constraints
------------------------------------------------------------

 num. of constraints = 50
 dim. of socp   var  = 1051,   num. of socp blk  = 501
*************************************************************************
   SDPT3: Infeasible path-following algorithms
*************************************************************************
 version  predcorr  gam  expon  scale_data
    NT      1      0.000   1        0
it pstep dstep pinfeas dinfeas  gap      prim-obj      dual-obj    cputime
-------------------------------------------------------------------------
 0|0.000|0.000|9.5e-01|2.3e+01|1.8e+04| 7.142482e+02  0.000000e+00| 0:0:00| chol  1  1
 1|1.000|1.000|2.7e-08|9.6e-02|7.2e+02| 6.546504e+02  9.246345e-01| 0:0:00| chol  1  1
 2|0.909|1.000|1.5e-07|9.6e-03|6.4e+01| 6.512372e+01  1.712575e+00| 0:0:00| chol  1  1
 3|0.934|0.434|4.6e-08|5.8e-03|1.8e+01| 2.184844e+01  3.858190e+00| 0:0:00| chol  1  1
 4|0.617|0.405|2.3e-08|3.5e-03|9.6e+00| 1.417260e+01  4.579879e+00| 0:0:00| chol  1  1
 5|0.536|0.728|1.1e-08|9.6e-04|5.0e+00| 9.606566e+00  4.625332e+00| 0:0:00| chol  1  1
 6|0.960|0.903|4.7e-10|9.4e-05|3.7e-01| 5.359583e+00  4.988527e+00| 0:0:00| chol  1  1
 7|0.845|0.966|2.0e-10|3.3e-06|9.5e-02| 5.109067e+00  5.014546e+00| 0:0:00| chol  1  1
 8|0.780|0.478|7.7e-09|1.7e-06|3.6e-02| 5.054972e+00  5.019026e+00| 0:0:00| chol  1  2
 9|0.813|0.679|3.0e-09|5.6e-07|1.3e-02| 5.034443e+00  5.021832e+00| 0:0:00| chol  2  2
10|0.736|0.767|8.0e-10|1.3e-07|5.2e-03| 5.027999e+00  5.022809e+00| 0:0:00| chol  2  2
11|0.749|0.738|2.0e-10|3.4e-08|2.0e-03| 5.025198e+00  5.023155e+00| 0:0:00| chol  2  2
12|0.584|0.623|8.4e-11|1.3e-08|1.0e-03| 5.024263e+00  5.023252e+00| 0:0:00| chol  2  2
13|0.806|0.752|1.9e-11|3.2e-09|3.0e-04| 5.023601e+00  5.023305e+00| 0:0:00| chol  2  2
14|1.000|0.805|3.3e-11|6.3e-10|4.9e-05| 5.023373e+00  5.023324e+00| 0:0:00| chol  2  2
15|0.743|0.923|1.4e-11|5.4e-11|1.4e-05| 5.023343e+00  5.023329e+00| 0:0:00| chol  2  2
16|0.811|0.924|6.5e-11|6.9e-12|4.3e-06| 5.023334e+00  5.023330e+00| 0:0:00| chol  3  3
17|0.627|0.727|1.6e-09|6.2e-12|2.3e-06| 5.023332e+00  5.023330e+00| 0:0:00| chol  2  2
18|0.601|1.000|6.6e-10|6.4e-12|1.2e-06| 5.023331e+00  5.023330e+00| 0:0:00| chol  2  2
19|0.605|1.000|2.7e-10|9.6e-12|6.7e-07| 5.023331e+00  5.023330e+00| 0:0:00| chol  2  2
20|0.588|1.000|1.2e-10|1.4e-11|3.7e-07| 5.023330e+00  5.023330e+00| 0:0:00| chol  2  2
21|0.590|1.000|5.6e-11|2.2e-11|2.0e-07| 5.023330e+00  5.023330e+00| 0:0:00| chol  2  2
22|0.594|1.000|3.1e-11|1.1e-11|1.1e-07| 5.023330e+00  5.023330e+00| 0:0:00|
  stop: max(relative gap, infeasibilities) < 1.49e-08
-------------------------------------------------------------------------

 number of iterations   = 22
 primal objective value =  5.02332996e+00
 dual   objective value =  5.02332985e+00
 gap := trace(XZ)       = 1.12e-07
 relative gap           = 1.01e-08
 actual relative gap    = 1.01e-08
 rel. primal infeas     = 3.09e-11
 rel. dual   infeas     = 1.12e-11
 norm(X), norm(y), norm(Z) = 3.2e+00, 5.8e-01, 2.5e+01
 norm(A), norm(b), norm(C) = 1.6e+02, 1.9e+01, 2.3e+01
```

```
 Total CPU time (secs)  = 0.38
 CPU time per iteration = 0.02
 termination code       =  0
 DIMACS: 8.2e-11  0.0e+00  1.3e-10  0.0e+00  1.0e-08  1.0e-08
------------------------------------------------------------------
------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +5.02333

Testing Error: 0.144338
Support: 43
```

B.2.2 Medium-scale dataset

```
Calling SDPT3 4.0: 10501 variables, 500 equality constraints
------------------------------------------------------------


 num. of constraints = 500
 dim. of socp   var  = 10501,   num. of socp blk  = 5001
*****************************************************************
   SDPT3: Infeasible path-following algorithms
*****************************************************************
 version  predcorr  gam  expon  scale_data
    NT      1      0.000   1        0
it pstep dstep pinfeas dinfeas  gap      prim-obj      dual-obj    cputime
-----------------------------------------------------------------
 0|0.000|0.000|9.8e-01|7.2e+01|5.2e+05| 7.093451e+03  0.000000e+00| 0:0:02| chol  1  1
 1|1.000|0.996|1.0e-08|3.7e-01|9.5e+03| 6.888375e+03  7.852670e-01| 0:0:05| chol  1  1
 2|0.999|1.000|2.0e-08|9.9e-03|5.1e+02| 5.111069e+02  1.218548e+00| 0:0:10| chol  1  1
 3|0.908|0.410|5.7e-08|6.2e-03|1.5e+02| 1.538825e+02  4.537091e+00| 0:0:15| chol  1  1
 4|0.714|1.000|1.5e-08|9.9e-05|5.9e+01| 6.176855e+01  2.879723e+00| 0:0:20| chol  1  1
 5|0.986|0.908|2.2e-10|1.8e-05|1.1e+01| 1.607598e+01  4.876689e+00| 0:0:25| chol  1  1
 6|0.912|0.981|4.0e-11|1.3e-06|9.8e-01| 5.994558e+00  5.010259e+00| 0:0:30| chol  1  1
 7|0.987|0.525|6.2e-09|6.7e-07|2.5e-01| 5.312589e+00  5.059905e+00| 0:0:35| chol  1  1
 8|0.702|0.572|2.4e-09|2.9e-07|1.2e-01| 5.215894e+00  5.095218e+00| 0:0:39| chol  1  1
 9|0.738|0.553|2.0e-10|1.3e-07|5.1e-02| 5.156974e+00  5.105812e+00| 0:0:44| chol  1  2
10|0.694|0.597|6.3e-11|5.3e-08|2.2e-02| 5.133150e+00  5.110765e+00| 0:0:49| chol  1  1
11|0.644|0.670|4.8e-11|1.8e-08|1.1e-02| 5.123456e+00  5.112953e+00| 0:0:54| chol  2  2
12|0.687|0.655|2.0e-11|6.1e-09|4.7e-03| 5.118430e+00  5.113733e+00| 0:0:58| chol  2  2
13|0.704|0.670|2.3e-11|2.0e-09|2.0e-03| 5.116103e+00  5.114053e+00| 0:1:03| chol  2  2
14|0.663|0.725|3.1e-11|5.6e-10|1.0e-03| 5.115217e+00  5.114186e+00| 0:1:08| chol  2  2
15|0.779|0.678|9.0e-11|1.9e-10|3.8e-04| 5.114611e+00  5.114232e+00| 0:1:13| chol  2  2
16|0.868|0.752|2.2e-10|5.5e-11|1.1e-04| 5.114364e+00  5.114253e+00| 0:1:18| chol  2  2
17|0.795|0.810|4.3e-10|2.4e-11|3.7e-05| 5.114298e+00  5.114261e+00| 0:1:23| chol  2  3
18|0.607|0.787|3.5e-10|2.6e-11|2.0e-05| 5.114282e+00  5.114263e+00| 0:1:28| chol  2  3
19|0.595|1.000|1.8e-10|3.1e-11|1.0e-05| 5.114274e+00  5.114264e+00| 0:1:33| chol  3  3
20|0.597|1.000|2.0e-10|3.6e-11|5.7e-06| 5.114269e+00  5.114264e+00| 0:1:38| chol  3  3
21|0.601|1.000|1.9e-10|4.1e-11|3.1e-06| 5.114267e+00  5.114264e+00| 0:1:43| chol  3  3
22|0.605|1.000|1.5e-10|3.9e-11|1.7e-06| 5.114266e+00  5.114264e+00| 0:1:47| chol  2  3
23|0.607|1.000|1.2e-10|3.1e-11|9.0e-07| 5.114265e+00  5.114264e+00| 0:1:52| chol  3  3
24|0.610|1.000|1.1e-10|2.4e-11|4.8e-07| 5.114264e+00  5.114264e+00| 0:1:57| chol  3  3
25|0.612|1.000|9.7e-11|2.3e-11|2.6e-07| 5.114264e+00  5.114264e+00| 0:2:02| chol  2  3
26|0.613|1.000|8.6e-11|1.9e-11|1.4e-07| 5.114264e+00  5.114264e+00| 0:2:07|
  stop: max(relative gap, infeasibilities) < 1.49e-08
-----------------------------------------------------------------
```

```
 number of iterations   = 26
 primal objective value =  5.11426403e+00
 dual   objective value =  5.11426389e+00
 gap := trace(XZ)       = 1.38e-07
 relative gap           = 1.23e-08
 actual relative gap    = 1.23e-08
 rel. primal infeas     = 8.56e-11
 rel. dual   infeas     = 1.95e-11
 norm(X), norm(y), norm(Z) = 3.2e+00, 6.2e-01, 8.1e+01
 norm(A), norm(b), norm(C) = 1.6e+03, 5.2e+01, 7.2e+01
 Total CPU time (secs)  = 126.66
 CPU time per iteration = 4.87
 termination code       =  0
 DIMACS: 4.6e-10  0.0e+00  7.0e-10  0.0e+00  1.2e-08  1.2e-08
------------------------------------------------------------------------
------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +5.11426

Testing Error: 0.078289
Support: 342
```