Optimization HW1

Gan Haochen 1600017789

April 2020

1 Newton-type Optimization

Newton-type Optimization is defined as the iteration procedure:

$$B_k d_k = -g_k \tag{1}$$

where $d_k := x_{k+1} - x_k$, $g_k := \nabla f(x_k)$. And using line-search to find an α such that

$$x_{k+1} = x_k + \alpha d_k \tag{2}$$

For Damping Newton method, the matrix B_k is the Hessian of the objective function:

$$B_k = G_k := \nabla^2 f(x_k) \tag{3}$$

For quasi-Newton method, the inverse-matrix $H_k = B_k^{-1}$ is defined and the iteration procedure could be re-written as:

$$d_k = -H_k g_k \tag{4}$$

and for the Broyden methods, the iteration of H_k is:

$$H_{k+1}^{\phi} = H_{k+1}^{\text{DFP}} + \phi(H_{k+1}^{\text{BFGS}} - H_{k+1}^{\text{DFP}}) = H_{k+1}^{\text{DFP}} + \phi v_k v_k^T \tag{5}$$

where

$$v_k := (y_k^T H_k y_k)^{\frac{1}{2}}\left(\frac{s_k}{s_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k}\right)$$

2 Implementation

2.1 Stopping Criterion

In our implementation the stopping criterion is setted as:

$$||g_k||_2 < \epsilon(1 + |f_k|), \quad \epsilon = 1\text{e} - 8 \tag{6}$$

Notice due to the magnitude of the function f, which could be very large or very small in some cases, an absolute criterion is not proper because of the machine accuracy.

2.2 Numerical Adjustment

In the iteration procedure of Broyden and SR1 methods, the denominator could be very small and cause problem. When this happen, we set $H_k = I_n$ where I_n is the identity of $n \times n$ dimensions. The identity is also used as the initial value of H_k.

2.3 C++ Eigen Library Used

A library for C++, Eigen is used for the linear system calculation. And VectorXd class is also very useful in the program.

3 Results

BDx means Brown-Dennis function with x dimension; DIE(x) means Discrete integral equation function with x dimension; CP mean Combustion Propane function from P153.

Timing of certain jobs of the optimization

$Problem$	DN	BFGS	Broyden($\phi = 0.5$)	DFS	SR1
$BD4$	0.313	0.758	1.093	12.766	4.393
$BD10$	0.276	0.564	0.645	12.928	22.08
$BD20$	0.174	0.328	0.282	0.283	3.392
$BD30$	0.22	0.24	0.203	0.208	15.211
$BD40$	0.171	0.226	0.188	0.222	19.323
$BD50$	0.159	0.257	0.233	0.214	39.021
$DIE2$	0.046	0.131	0.064	0.05	0.086
$DIE10$	0.156	0.32	0.32	0.295	0.068
$DIE20$	0.27	0.503	0.502	0.401	0.098
$DIE30$	0.308	0.814	0.784	0.682	0.177
$DIE40$	0.495	1.006	0.784	1.001	0.256
$DIE50$	0.396	0.563	0.784	0.545	0.363
$CP(P153)$	/	0.298	2.369	23.621	/

3.1 Analysis

Note that at least in our implementation Damping Newton is almost always better than Broyden-type method. In BD problem the performance of SR1 method is not so satisfying. But in DIE problem the SR1 method works well.

And note that DFS method just failed for small scale BD problem, which could because of the denominator is still too small in the formula. And during the experiment we have observed that the line search is hard for Broyden-type method when close to the optimal. And at the same time Damping Newton method does not evoke too many line researches.

iteration number

$Problem$	DN	BFGS	Broyden($\phi = 0.5$)	DFS	SR1
$BD4$	20	76	106	/	/
$BD10$	15	50	70	/	1999
$BD20$	9	24	25	23 325	
$BD30$	7	13	12	13	838
$BD40$	7	11	11	11	750
$BD50$	7	10	11	9	1573
$DIE2$	4	14	13	11	13
$DIE10$	4	14	14	14	12
$DIE20$	4	14	14	13	11
$DIE30$	4	14	14	13	11
$DIE40$	4	14	14	14	11
$DIE50$	4	14	14	14	11
$CP(P153)$	/	70	439	1981	/

function evoking number

$Problem$	DN	BFGS	Broyden($\phi = 0.5$)	DFS	SR1
$BD4$	61	244	579	/	/
$BD10$	46	169	397	/	20203
$BD20$	28	94	191	282	2618
$BD30$	22	67	131	197	8818
$BD40$	23	68	136	2042	9375
$BD50$	23	727	147	215	23596
$DIE2$	13	15	30	45	53
$DIE10$	13	55	110	165	49
$DIE20$	137	55	110	161	49
$DIE30$	138	55	110	161	45
$DIE40$	13	55	110	165	45
$DIE50$	13	55	110	165	45
$CP(P153)$	/	287	2495	22353	/

And for the DIE problem we noticed that Damping Newton method just arrived 10^{-34} order, while other methods cannot. This is probably because of the machine accuracy and because Newton method could reach a very close point at once without being influenced by machine accuracy.

iteration number

$Problem$	DN	BFGS	Broyden($\phi = 0.5$)	DFS	SR1
$BD4$	20	76	106	/	/
$BD10$	15	50	70	/	1999
$BD20$	9	24	25	23	325
$BD30$	7	13	12	13	838
$BD40$	7	11	11	11	750
$BD50$	7	10	11	9	1573
$DIE2$	4	14	13	11	13
$DIE10$	4	14	14	14	12
$DIE20$	4	14	14	13	11
$DIE30$	4	14	14	13	11
$DIE40$	4	14	14	14	11
$DIE50$	4	14	14	14	11
$CP(P153)$	/	70	439	1981	/

Optimal

$Problem$	DN	BFGS	Broyden($\phi = 0.5$)	DFS	SR1
$BD4$	$1.05E-05$	$1.05E-05$	$1.05E-05$	$1.08E-05$	$1.05E-05$
$BD10$	1.44323	1.44323	1.44323	1.44323	1.44323
$BD20$	85822.2	85822.2	85822.2	85822.2	85822.2
$BD30$	$9.77E+08$	$9.77E+08$	$9.77E+08$	$9.77E+08$	$9.77E+08$
$BD40$	$5.86E+12$	$5.86E+12$	$5.86E+12$	$5.86E+12$	$5.86E+12$
$BD50$	$2.67E+16$	$2.67E+16$	$2.67E+16$	$2.67E+16$	$2.67E+16$
$DIE2$	$3.01E-34$	$6.83E-20$	$2.17E-20$	$1.11E-21$	$1.77E-17$
$DIE10$	$3.33E-35$	$1.70E-17$	$1.70E-17$	$1.70E-17$	$1.70E-17$
$DIE20$	$4.99E-35$	$2.31542E-17$	$2.31542E-17$	$2.31542E-17$	$2.31542E-17$
$DIE30$	$1.09E-34$	$1.64E-17$	$1.60E-17$	$2.37E-17$	$2.12E-17$
$DIE40$	$1.20E-34$	$1.20E-34$	$1.69E-17$	$1.64E-17$	$1.58E-17$
$DIE50$	$1.48E-34$	$1.72E-17$	$1.68E-17$	$1.64E-17$	$2.26E-17$
$CP(P153)$	/	$4.30E-18$	$2.55E-17$	$1.50E-01$	/