

# STAT 31020: Homework 4

Caleb Derrickson

January 31, 2024

## Contents

<b>1</b>	<b>Problem 1</b>	<b>2</b>
1	Problem 1, part 1 . . . . .	2
2	Problem 1, part 2 . . . . .	3
3	Problem 1, part 3 . . . . .	4
4	Problem 1, part 4 . . . . .	5
<b>2</b>	<b>Problem 2</b>	<b>6</b>
<b>3</b>	<b>Problem 3</b>	<b>8</b>
1	Problem 3, part a . . . . .	8
2	Problem 3, part b . . . . .	10
3	Problem 3, part c . . . . .	11
4	Problem 3, part d . . . . .	12
5	Problem 3, part e . . . . .	12
6	Problem 3, part f . . . . .	13

## Problem 1

We aim to understand both trust-region approaches and sparsity of matrices.

### Problem 1, part 1

For the Rosenbrock function of order  $n$  we used in last homework,

$f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$  implement a function to return the sparse Hessian.

---

#### **Solution:**

I have just added the condition that the returned hessian is sparse. I passed the dense matrix to the `sparse()` function provided by Matlab.

## Problem 1, part 2

We will experiment with sparse and dense linear algebra. Let  $b \in \mathbb{R}^n$  be a vector of ones of dimension  $n$ . Compute the Hessian  $H$  at the point of all entries being equal to 2. Solve the equation  $Hx = b$  and time it for  $H$  being sparse format and denote the time by  $a(n)$  with dense format and denote the time by  $b(n)$  using the backslash operator. Plot  $a(n)$  and  $b(n)$  as functions of  $n$ . what seems to be the rough behavior of  $a(n)$  and  $b(n)$ ?

---

### Solution:

The plot is given as Figure 1. I tried fitting polynomials to both plots and I got that the dense matrix time scales roughly by  $n^2$  (the coefficient of  $n^3$  is of the order  $10^{-11}$ ; for  $n^2$ ,  $10^{-8}$ ), while the sparse matrix scales roughly by  $n$  (the coefficient of  $n^2$  is of order  $10^{-10}$ ; for  $n$ ,  $10^{-7}$ ). I have been told that I will not see the  $n^3$  dominance come into play for such small dimension values in the case of the dense matrix. The way I was constructing my Hessian, I had a nested for loop which really dropped computation time.

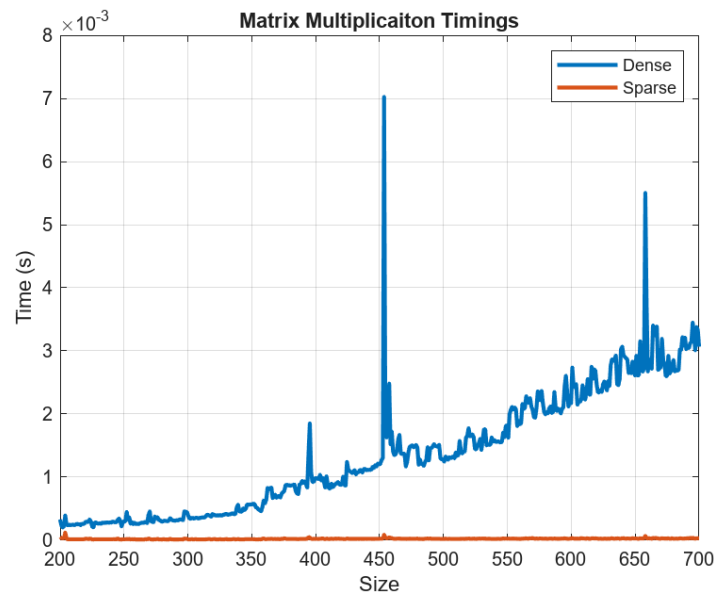


Figure 1: Solving  $Ax = b$  for  $A$  dense vs sparse and the given Hessian.

## Problem 1, part 3

Write a program that implements the dogleg trust-region approach. Experiment with the parameters or by designing your own rules. Report the total number of linear systems solved and the total number of function evaluations.

---

### Solution:

I have (hopefully) implemented the Dogleg method correctly. Using max 500 steps, initial point  $x_0 = [7 \ 7 \ 7 \ 7 \ 7]$ , and epsilon = 0.8<sup>1</sup>, I get the normed gradient plot below. Fixing all other parameters (tolerance = 0.75, eta = 0.9 \* tolerance), I found that the convergence of the sequence to a respectable value is highly sensitive to the value of  $\hat{\Delta}$ . The reported plot is the best value I could find. The number of iterations in this plot is 103, with the number of function evaluations being 309 (two for calculating  $\rho_k$ , one in Dogleg proper) and the number of linear systems solved being 103 (only one per loop, in finding the newton step). Note that I am defining the number of function evaluations as the number of times the literal state function is being evaluated: not counting the Hessian or the gradient.

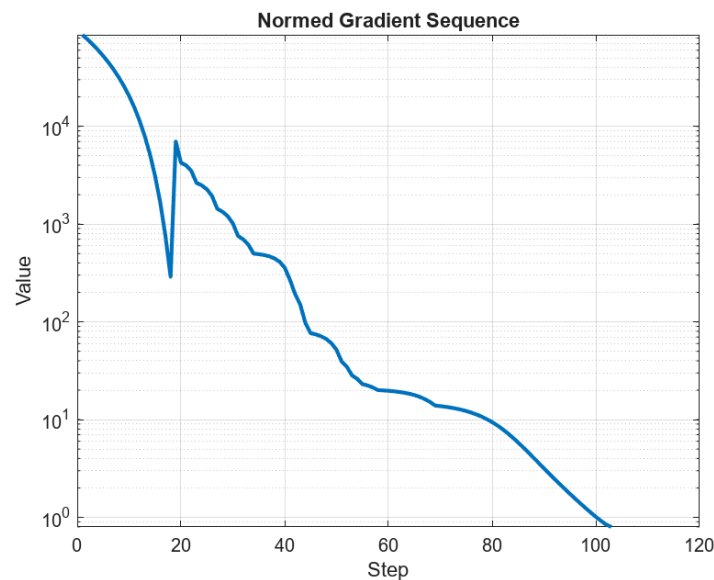


Figure 2: Dog Leg Method for initial condition described above, with  $\hat{\Delta} = 4.25$ .

---

<sup>1</sup>I know it's bad, I don't know what's wrong in my code and I have spent too much time arguing with my code, running it line by line over 200 iterations to see where it went wrong. I had the same issue for NMHM as well in last homework.

## Problem 1, part 4

Compare NMHM (with dense linear algebra) from the previous homework and dogleg in terms of (a) number of systems solved (b) total computation time of their linear systems by starting them at the point of all 2's using the same stopping criterion for the gradient for both. Plot the two as a function of problem size for some ranges of  $n$ . Note that the  $n$  range need not be the same for both as the timings may be different. Discuss your findings and explain why the results may behave as you observe.

---

### Solution:

I have included the two relevant plots below. For the NMHM plot, I know there is something wrong with the normed gradient converging to zero. I can say that after a weekend's worth of debugging, that it *does* converge, just not to zero. I honestly don't know why. We can see that the average time to solve the linear system is greater for Dogleg, but there are overall fewer linear systems calculated. To somehow rationalize my findings, I would say that the say that sparse matrices are stored in memory isn't contiguous, unlike the case for dense matrices. And as such, for lower dimensions as what I have here, we see this difference. PLEASE BE IN THE MATLABCODE FOLDER WHEN RUNNING ALL CODE. IF YOU ARE NOT, YOU WILL GET AN ERROR. I wanted to separate the two methods via folders, so I change the current directory when calling the two methods.

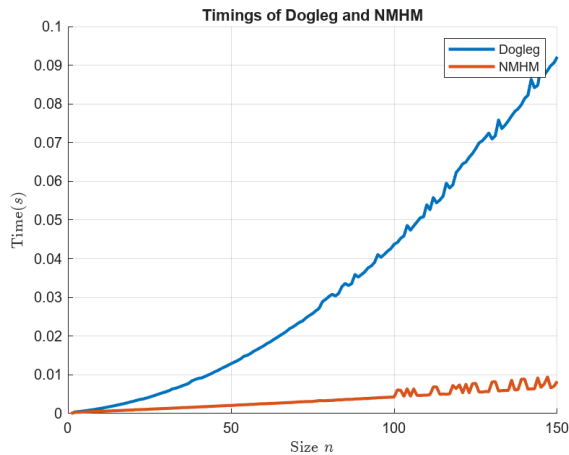


Figure 3: The average time to solve the linear system for NMHM and Dogleg methods.

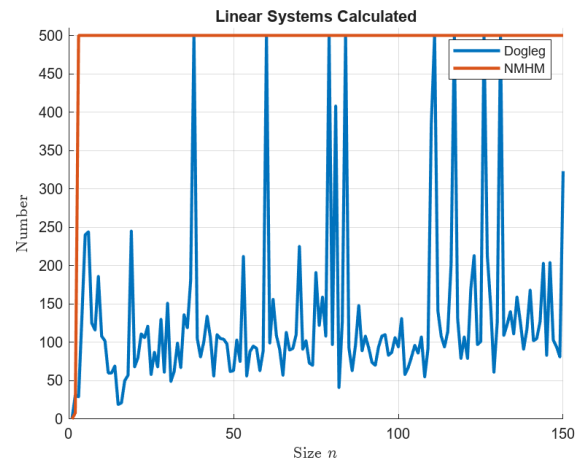


Figure 4: The number of linear systems calculated up to 150 dimensions.

## Problem 2

This problem we will carry out what I mentioned in class when I realized my insight was incorrect. Assume that you try to solve the trust-region problem approximately, as I did in class with the dogleg. In other words you aim to solve the problem

$$\min_p := f + g^\top p + \frac{1}{2} p^\top B p, \quad \|p\| \leq \Delta.$$

Assume now that (a)  $B$  is invertible and has at least one negative eigenvalue, (b)  $p^N$  is a solution of the Newton system  $Bp^N = -g$  and (c)  $\|p^N\| < \Delta$  (that is, the Newton point is interior to the trust region, the situation I was implementing in class). Prove now that there exists a direction  $p^D$  such that  $m(p^N + \varepsilon p^D) < m(p^N)$  for  $\varepsilon$  sufficiently small. In other words,  $p^N$  cannot be a solution of the trust region method even if it is a solution of the Newton equation.

### Solution:

We will begin by expanding  $m(p^N + \varepsilon p^D)$

$$\begin{aligned} m(p^N + \varepsilon p^D) &= f + g^\top (p^N + \varepsilon p^D) + \frac{1}{2} (p^N + \varepsilon p^D)^\top B (p^N + \varepsilon p^D) \\ &= f + g^\top p^N + \varepsilon g^\top p^D + \frac{1}{2} \left[ (p^N)^\top B p^N + \varepsilon (p^N)^\top B p^D + \varepsilon (p^D)^\top B p^N + \varepsilon^2 (p^D)^\top B p^D \right] \\ &= f + g^\top p^N + \frac{1}{2} (p^N)^\top B p^N + \varepsilon g^\top p^D + \frac{1}{2} \left[ \varepsilon ((p^N)^\top B p^D + (p^D)^\top B p^N) + \varepsilon^2 (p^D)^\top B p^D \right] \\ &= m(p^N) + \varepsilon \left( g^\top p^D + \frac{1}{2} (p^N)^\top B p^D + \frac{1}{2} (p^D)^\top B p^N + \frac{\varepsilon}{2} (p^D)^\top B p^D \right) \end{aligned}$$

By the calculations above, we just need to show that the quantity in parenthesis can be negative for some choice of search direction  $p^D$ . If we show this, then we will show that  $p^N$  cannot be a solution. Note that from (b) we have  $Bp^N = -g$ . Plugging this into the quantity in parenthesis, we have

$$-(p^N)^\top B^\top p^D + \frac{1}{2} (p^N)^\top B p^D + \frac{1}{2} (p^D)^\top B p^N + \frac{\varepsilon}{2} (p^D)^\top B p^D$$

By transposition, we can rewrite every suitable instance of  $p^D$  being pre-multiplied by  $B$ . Then,

$$-(p^N)^\top B p^D + \frac{1}{2} (p^N)^\top B p^D + \frac{1}{2} (B^\top p^D)^\top p^N + \frac{\varepsilon}{2} (p^D)^\top B p^D$$

Since we have that  $B$  has one negative eigenvalue, suppose that  $p^D$  is the eigenvector associated with this eigenvalue. We can then write  $Bp^D = -\lambda p^D$ , for  $\lambda > 0$ . By the properties of linear algebra,  $\lambda$  is also an eigenvalue of  $B^\top$  since we are over the real numbers. Thus, the following can be written:

---


$$- (p^N)^\top B p^D + \frac{1}{2} (p^N)^\top B p^D + \frac{1}{2} (B^\top p^D)^\top p^N + \frac{\varepsilon}{2} (p^D)^\top B p^D \quad (\text{Given.})$$

$$- (p^N)^\top (-\lambda) p^D + \frac{1}{2} (p^N)^\top (-\lambda) p^D + \frac{1}{2} (-\lambda p^D)^\top p^N + \frac{\varepsilon}{2} (p^D)^\top (-\lambda) p^D \quad (B p^D = -\lambda p^D.)$$

$$\lambda (p^N)^\top p^D - \frac{\lambda}{2} (p^N)^\top p^D - \frac{\lambda}{2} (p^D)^\top p^N - \frac{\lambda \varepsilon}{2} (p^D)^\top p^D \quad (\text{Rearranging.})$$

$$\lambda (p^N)^\top p^D - \frac{\lambda}{2} (p^N)^\top p^D - \frac{\lambda}{2} (p^N)^\top p^D - \frac{\lambda \varepsilon}{2} (p^D)^\top p^D \quad (\text{Inner product symmetry.})$$

$$= -\frac{\lambda \varepsilon}{2} (p^D)^\top p^D \quad (\text{Cancelling.})$$

$$= -\frac{\lambda \varepsilon}{2} \|p^D\|^2 \quad (\text{Definition.})$$


---

Since  $\lambda, \varepsilon > 0$ , we have that the above term is overall negative. Therefore, an additional direction  $p^D$  has been obtained which minimizes the approximate solution.

## Problem 3

We will show that steepest descent with backtracking has asymptotic linear convergence. Let  $x_k$  be a sequence produced by steepest descent with line search and backtracking when applied to the nonlinear optimization problem  $\min_x f(x)$ , where the function  $f(x)$  is three times continuously differentiable. This is the algorithm you implemented in the second homework except you will apply it to a general function. In particular, we will assume that backtracking is initialized at a fixed stepsize  $\bar{\alpha}$  as we did in problem 2 in homework 2. Assume that  $x_k \rightarrow x^*$  where  $x^*$  satisfies the second order sufficient conditions.

### Problem 3, part a

Prove that there exists a neighborhood of  $x^*$  and parameters  $0 < c_1 < c_2$  such that  $c_1 \|\nabla f(x)\|^2 \leq f(x) - f(x^*) \leq c_2 \|\nabla f(x)\|^2$  for any  $x$  within the neighborhood  $N(x^*)$ .

---

#### Solution:

The first inequality can be shown as follows. We first note that the neighborhood  $N(x^*)$  can be chosen around  $x^*$  such that  $f(x \in N(x^*))$  is convex, as well as the gradient  $\nabla f(x)$  is Lipschitz. Note that [1] gives<sup>2</sup>.

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{1}{2L} \|\nabla f(y) - \nabla f(x)\|^2, \quad \forall x, y.$$

This result is directly applicable when we restrict  $x, y \in N(x^*)$ , so that we can assure convexity of  $f$  and Lipschitzness of  $\nabla f$ . Substituting  $y \mapsto x$  and  $x \mapsto x^*$  gives,

$$f(x) \geq f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2L} \|\nabla f(x) - \nabla f(x^*)\|^2.$$

By the second order sufficient conditions, we have that  $\nabla f(x^*) = 0$ , so we get

$$\frac{1}{2L} \|\nabla f(x)\|^2 \leq f(x) - f(x^*).$$

So we have  $c_1 = \frac{1}{2L}$ . To find the second inequality, we can approximate the function  $f(x)$  for some  $x \in N(x^*)$ ,

$$f(x) = f(x^*) + \nabla f(x^*)^\top (x - x^*) + \frac{1}{2} (x - x^*)^\top \nabla^2 f(x^*) (x - x^*).$$

By the second order sufficient conditions we have  $\nabla f(x^*) = 0$ , so with rearranging, we have

$$f(x) - f(x^*) = \frac{1}{2} (x - x^*)^\top \nabla^2 f(x^*) (x - x^*)$$

---

<sup>2</sup>This result is [5] on page 4. I have two choices concerning this: I can either give my own, sloppier version of the proof; or I can cite the reference itself. I feel it is more transparent that I cite it.



The right hand side can be bounded by Cauchy Schwartz to give

$$f(x) - f(x^*) \leq \|x - x^*\|^2 \|\nabla^2 f(x^*)\|.$$

By problem 3 part a of Homework 2 we found

$$c' \|x - x^*\| \leq \|\nabla f(x)\|,$$

for some  $c' > 0$ . We can substitute this into our inequality to get

$$f(x) - f(x^*) \leq \frac{1}{2(c')^2} \|\nabla^2 f(x^*)\| \|\nabla f(x)\|^2.$$

Therefore, setting  $c_2 = \frac{1}{2(c')^2} \|\nabla^2 f(x^*)\|$ , we have found the following chain of inequalities to be found.

### Problem 3, part b

In the case where  $f(x)$  is quadratic,  $f(x) = f + g^\top x + \frac{1}{2}x^\top Bx$ ,  $B \succ 0$  find the sharpest values of  $c_1, c_2$  as a function of the eigenvalues of  $B$ . Prove that they are the sharpest by finding  $x \neq x^*$ , where the inequalities above are, in effect, equalities.

---

#### Solution:

For the case of the quadratic, we can take some shortcuts in the calculation of  $c_1$  and  $c_2$  in the previous part. Define  $\psi(x) = f(x) - f$ . We note that analyzing  $\psi$  over  $f(x)$  will return equivalent bounds. Note that

$$\psi(x) = \frac{1}{2} \langle x - x^* | B(x - x^*) \rangle \leq \frac{\|B\|}{2} \|x - x^*\|^2 = \frac{\lambda_{\max}}{2} \|x - x^*\|^2.$$

From the previous homework, we found that  $\|x - x^*\| \leq \frac{1}{\lambda_{\max}^2} \|\nabla f(x)\|^2$ , so substituting this in, we find

$$\psi(x) = f(x) - f \leq \frac{1}{2\lambda_{\max}} \|\nabla f(x)\|^2.$$

This then settles the upper bound. We can then find the lower bound in the following manner:

DO THIS ONE.

We now have  $c_1 = \frac{1}{2\lambda_{\min}}$ ,  $c_2 = \frac{1}{2\lambda_{\max}}$ .

### Problem 3, part c

Use point (a) to show that one our iteration stays in that neighborhood of  $x^*$ , we will have

$$f(x^{k+1}) - f(x^*) \leq (1 - \eta_k)(f(x^k) - f(x^*))$$

for some  $\eta_k > 0$ . I suggest you to start form the analysis you did at Problem 2, homework 2 and obtain  $\eta_k$  from there. Note that via the iteration, we can rewrite  $\|p_k\|$  in terms of  $\|\nabla f_k\|$ . This is found from  $x_{k+1} = x_k + \alpha_k p_k$ . which can be rewritten as  $\|p_k\| \geq \|x_k - x^*\|$ . Then we can apply part a from last homework to this to get an inequality to  $\|\nabla f_k\|$

---

#### Solution:

Note that the given inequality can be rewritten in the following form:

$$f(x^{k+1}) \leq f(x^k) + \eta_k(f(x^*) - f(x^k))$$

From Homework 2, part 2, we were analyzing a sequence which was of the form  $x^{k+1} = \alpha^k + p^k$ . We will start in a similar manner as in that problem. Note I will be shorthanding  $f(x^i) = f_i$ .

---

$$f_{k+1} \leq f_k + c_1 \alpha_k \nabla f_k^\top p_k \quad (\text{Start.})$$

$$\leq f_k + c_1 \rho \bar{\alpha} \nabla f_k^\top p_k \quad (\text{Backtracking bound.})$$

$$\leq f_k - c_1 \rho \bar{\alpha} \|\nabla f_k^\top p_k\| \quad (\text{Norm of negative value.})$$

$$\leq f_k - c_1 \rho \bar{\alpha} \|\nabla f_k\| \|p_k\| \quad (\text{Cauchy Schwartz.})$$

$$\leq f_k - c_1 \rho \bar{\alpha} \|\nabla f_k\|^2 \quad (\text{Iteration.})$$

$$\leq f_k - c_1 \rho \bar{\alpha} (2\lambda_{\max})(f(x_k) - f(x^*)) \quad (\text{From part a.})$$

$$\implies \eta_k = 2c_1 \rho \bar{\alpha} \lambda_{\max} \quad (\text{Comparing.})$$

---

Note that  $\lambda_{\max}$  is the max eigenvalue of  $B_k$ , which is  $\eta$ 's dependence on  $k$ .

### Problem 3, part d

Show now that  $\eta_k$  is uniformly bounded below away from zero, and concluded that  $f(x^k) - f(x^*)$  converges Q-linearly to zero.

---

#### Solution:

Note that  $B_k$  is an approximation to the Hessian  $\nabla^2 f(x_k)$ . As  $k \rightarrow \infty$ , we have that  $\|B_k\| \rightarrow \|\nabla^2 f(x^*)\|$ . Since  $x^*$  satisfies the second order sufficient conditions, we have that  $\nabla^2(f(x^*)) \succ 0$ , which implies that  $\lambda_{\max}$  is bounded away from zero. Due to this reason, taking norms on each side of the inequality in part c, we get

$$\|f_{k+1} - f(x^*)\| \leq |1 - \eta_k| \|f_k - f(x^*)\|.$$

Dividing through by  $\|f_k - f(x^*)\|$  tells us that the ratio between function evaluations is bounded by  $|1 - \eta_k|$ , which we showed was bounded away from zero. Therefore,  $f(x^k) - f(x^*)$  converges Q-linearly to zero.

### Problem 3, part e

Show that the sequence  $\|x_k - x^*\|$ ,  $\|\nabla f(x^k)\|$  converge R-linearly to zero.

---

#### Solution:

To show these two sequences converge R-linearly to zero, we just need to show that there is some other sequence  $r_k$  which bounds the sequence and converges Q-linearly to zero. Note that by part a,  $\|\nabla f(x)\|^2$  is bounded by  $\frac{1}{c_1}(f(x) - f(x^*))$ . Since we showed the sequence  $f(x_k) - f(x^*)$  converges Q-linearly to zero, and it bounds  $\|\nabla f(x)\|^2$  for any  $x$  in our neighborhood of  $x^*$ , then  $\|\nabla f(x_k)\|$  converges R-linearly to zero. Doing the same procedure with the relationship we found between  $\|x - x^*\|$  and  $\|\nabla f(x)\|$  in Homework 2, Problem 3, we have that  $\|x^k - x^*\|$  converges R-linearly to zero. Note that we are free to apply this result from problem 3, since we satisfy the same conditions.

### Problem 3, part f

Is it possible for the convergence rate to be quadratic?

---

#### **Solution:**

Note that we are only using information about the gradient, not the Hessian. Only Newton's method converges quadratically, which requires the exact hessian of the function to work. Thus, we are only guaranteed at most superlinear convergence.