AE6310: Optimization for the Design of Engineered Systems
Quiz 1
February 11th, 2020

Briefly answer the following questions on the paper provided. Organize your work and be careful to properly answer all parts of each question.
This quiz is closed book. The length of the quiz is 40 minutes.

1. (25 points) For each part, sketch the contours of the quadratic function

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c$$

with the properties described below. Indicate any critical points and identify whether they are local minimum, maximum or saddle points. Note that the answer is not unique.

   (a) The matrix A is positive definite and the eigenvalues of A are not equal
   (b) The matrix A has positive and negative eigenvalues
   (c) The vector $b = 0$ and A is positive semi-definite

2. (25 points) Find and characterize the critical points of the following function. Use the unconstrained optimality theory to justify your answer.

$$f(x_1, x_2) = \frac{1}{2}x_1^4 - x_1^2 + x_2^2 - 3$$

3. (25 points) The following question deals with line searches and line search criteria. Create different sketches for part (b) and (c).

   (a) State the strong Wolfe conditions for the merit function $\phi(\alpha)$.
   (b) Sketch a valid merit function that has an interval that satisfies the curvature condition, but not the sufficient decrease condition.
   (c) Sketch a new merit function and the sufficient decrease condition. Illustrate on the merit function, the α points a backtracking line search method would evaluate if the first line search step failed, but the second one passed. Describe briefly how the backtracking line search algorithm works.

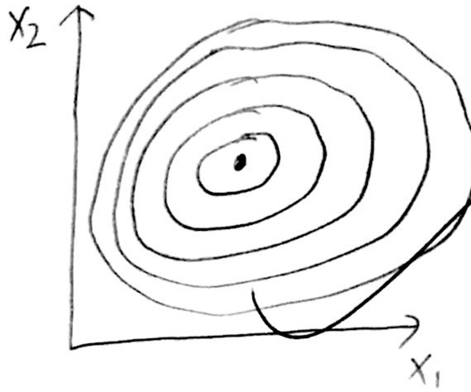
4. (25 points) You are given a quadratic function with a Hessian matrix that has the following properties:

$$\Lambda = \begin{bmatrix} 24 & 0 \\ 0 & 4 \end{bmatrix} \qquad Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad b = 0$$

   (a) Sketch the quadratic function
   (b) Re-draw the sketch and illustrate the convergence behavior of the steepest descent method using exact line searches from the point $x_1 = 2$, $x_2 = 1$.
   (c) Re-draw the sketch and illustrate the convergence behavior of the conjugate gradient method using exact line searches from the point $x_1 = 2$, $x_2 = 1$.
   (d) Describe why you expect to see a difference in performance between these two methods.
   (e) Describe any differences you would observe if you used inexact line searches.

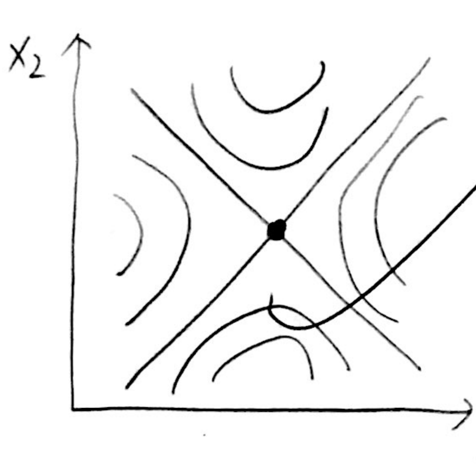
AE 6310 Quiz 1

1. (a)



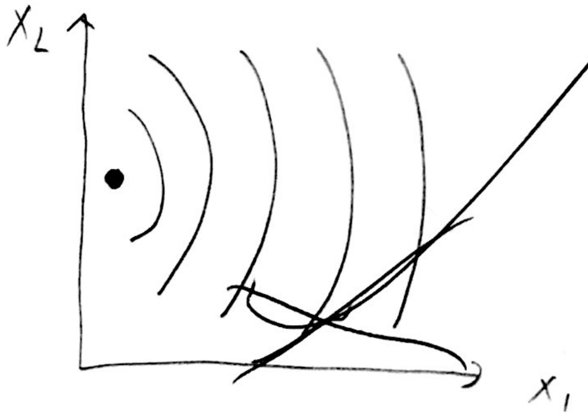
local minimu

(b)

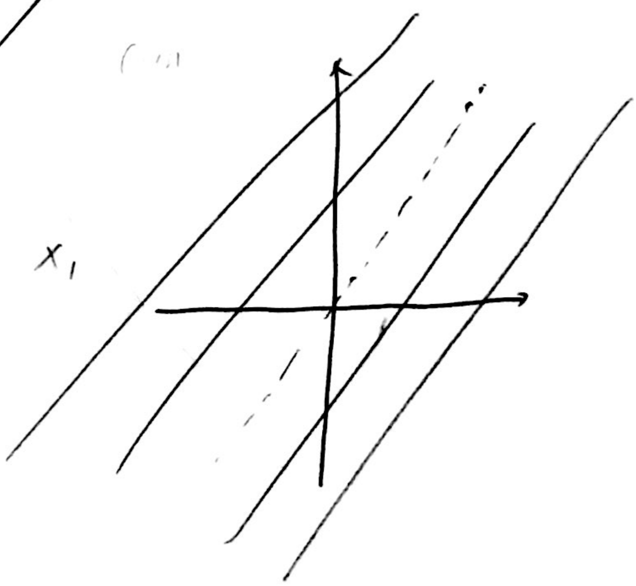


Saddle point

(c)



local maximum



- 8

2) $f(x_1, x_2) = \frac{1}{2} x_1^4 - x_1^2 + x_2^2 - 3$

$$\nabla f = \begin{bmatrix} 2x_1^3 - 2x_1 \\ 2x_2 \end{bmatrix}$$

$$H = \begin{bmatrix} 6x_1^2 - 2 & 0 \\ 0 & 2 \end{bmatrix}$$

To identify critical points, set $\nabla f = 0$ and evaluate x!

$$\left.\begin{aligned} 2x_1^3 - 2x_1 &= 0 \\ 2x_2 &= 0 \end{aligned}\right\} \Rightarrow \begin{aligned} x_1(2x_1^2 - 2) &= 0 \\ 2x_2 &= 0 \end{aligned}$$

Finding, $(-1, 0)$, $(0, 0)$, $(1, 0)$

For $\vec{x} = (-1, 0)$, $H = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$ → local minimum since positive definite

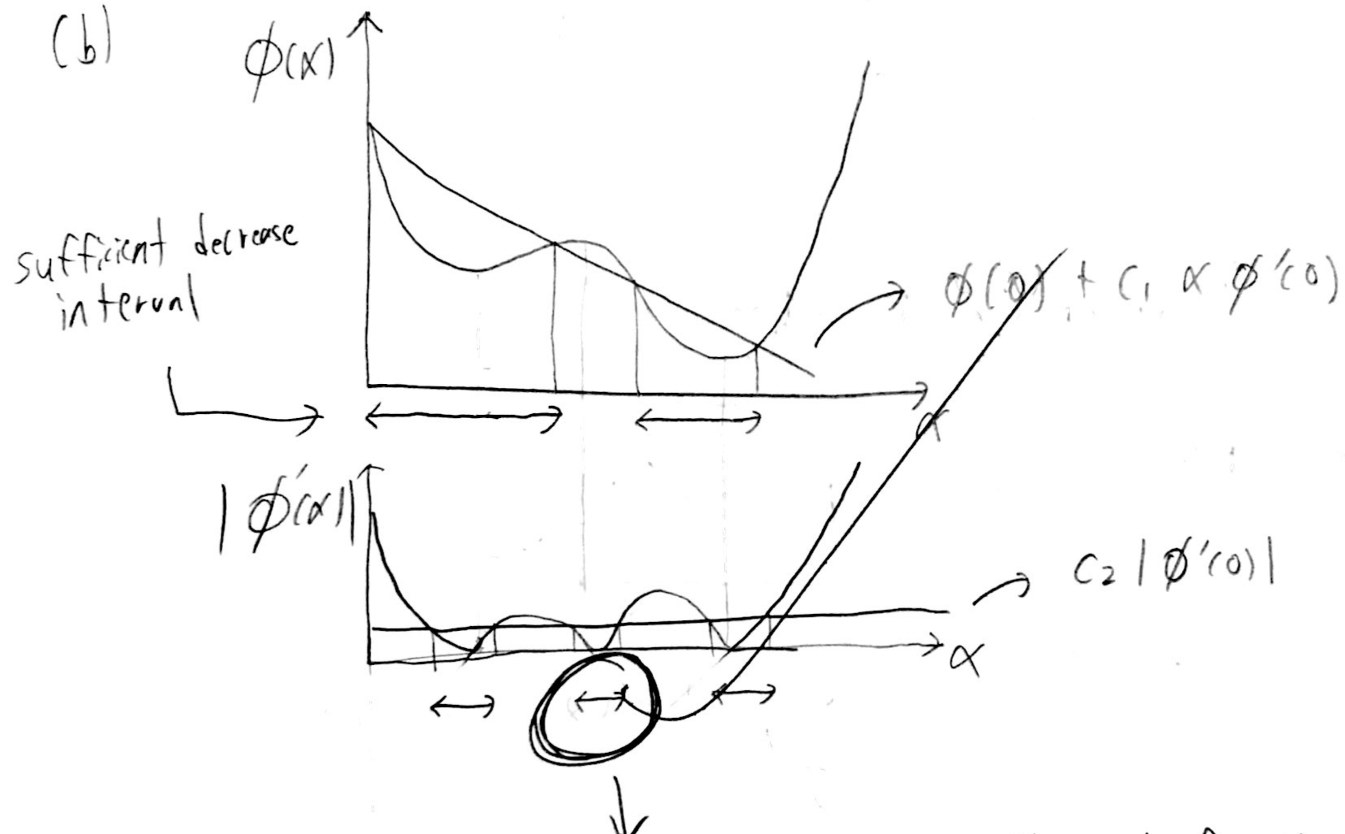
$\vec{x} = (0, 0)$, $H = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$ → saddle point since indefinite

$\vec{x} = (1, 0)$, $H = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$ → local minimum since positive definite

(a) $\phi(\alpha) \le \phi(0) + c_1 \alpha \phi'(0)$ $\quad c_1 \approx 10^{-3} - 10^{-4} \in (0,1)$

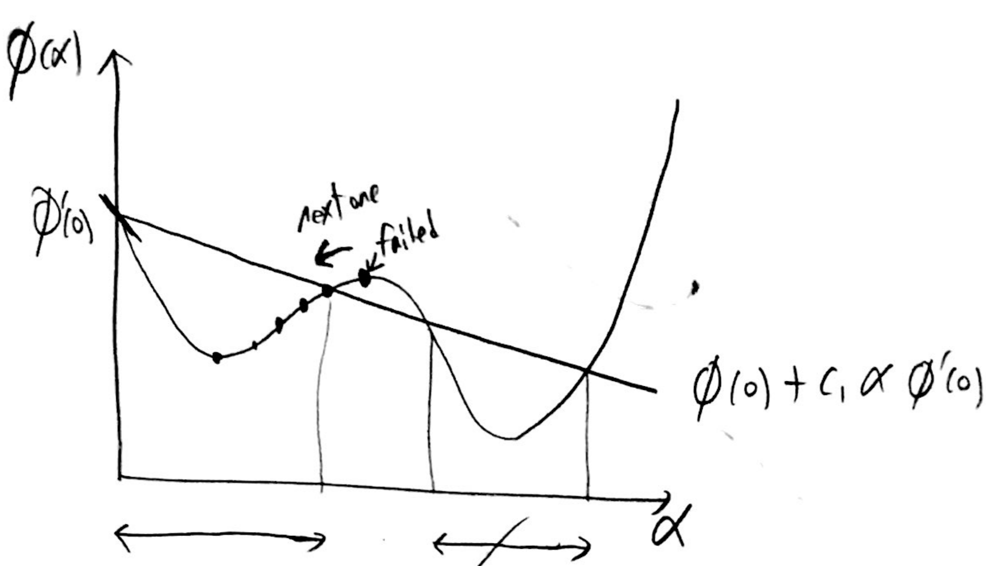
$|\phi'(\alpha)| \le c_2 |\phi'(0)|$ $\quad c_2 \in (c_1, 1)$ $\approx 0.1 - 0.9$

(b)



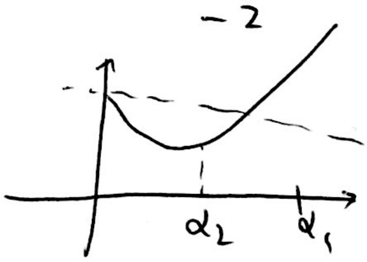
This is where the valid merit function satisfies the curvature condition, but not the sufficient decrease condition.

(c)



If the first linesearch step failed, but the second one passed, the algorithm would back up and continue search.

more explain.



$\alpha_2 = \tau \alpha_1 \quad \tau \in (0,1)$

$\Lambda = \begin{bmatrix} 24 & 0 \\ 0 & 4 \end{bmatrix} \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad b = 0$

(a) Since $Q^T \Lambda Q = A$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 24 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 24 & 0 \\ 0 & 4 \end{bmatrix} = A$$

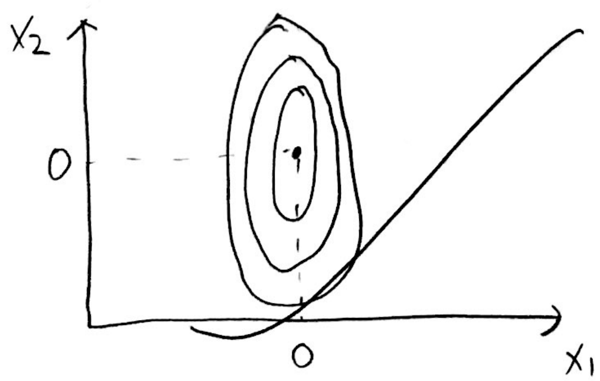
$$\Rightarrow f(x_1, x_2) = \frac{1}{2} x^T A x$$

$$= \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 24 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

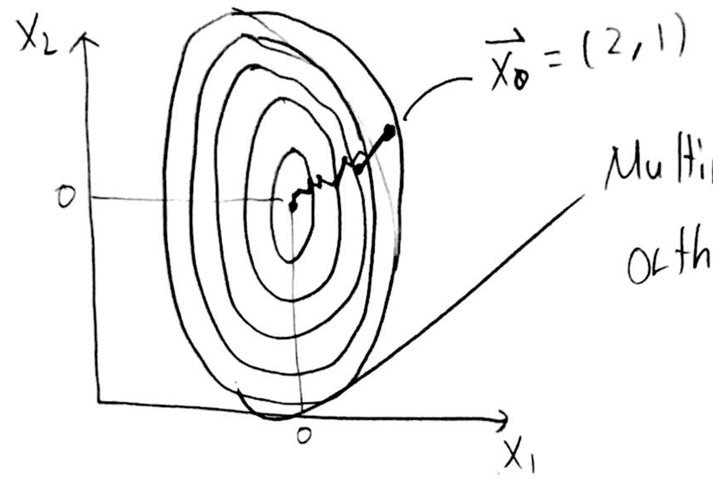
$$= \frac{1}{2} \begin{bmatrix} 24x_1 & 4x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \frac{1}{2} (24x_1^2 + 4x_2^2)$$

$$\boxed{\therefore f(x_1, x_2) = 12x_1^2 + 2x_2^2}$$

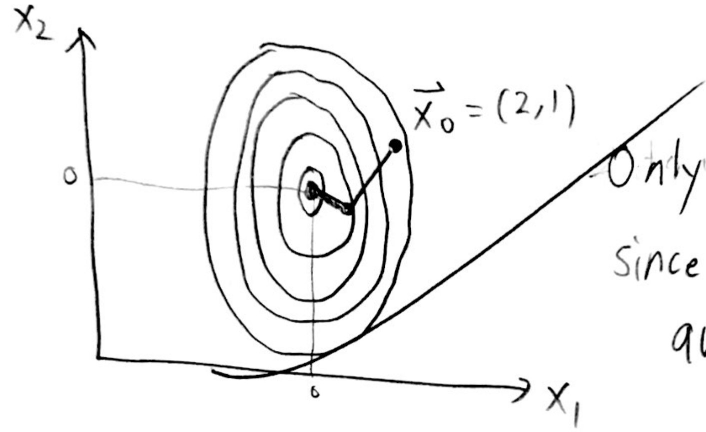


(b)



Multiple iterations (slow)
orthonormal directions

(c)



Only two iterations
since the function is
quadratic!

(d) The steepest descent method approaches to the optimal solution in an orthonormal manner and therefore slow. On the other hand, the conjugate gradient method requires at most n steps for a quadratic function where n is derived from $[p_1, p_2, \cdots p_n]$ directions. In this case, since $n = 2$, we only require 2 steps!

(e) For inexact line searches, we expect more iterations.
In this case, the steepest descent method may not have perfectly orthonormal directions each iteration due to approximation.
Approximation also holds true for conjugate gradient method.

C. G also needs more iterations. -1

4