On Studying Hessian Matrix with Applications

Adel Mohammed Hassan Rizak Al-Rammahi
Assistant Professor ,PH.D, math.Dep.,math. & comp.Coll.,Kufa Un, Iraq
E-mail: adelr@naharpost.com

Abstract

To study the properties of definite and Hessian matrices and using it in finding the critical points of quadratic forms .

On the other hand, the applications of Hessian matrix are introduced in static springs system problem .

حول دراسة مصفوفة هيسيان مع التطبيقات

د.عـادل مـحـمـد حسـن رزاق الرماحـــي
أستـاذ مسـاعد ـ قسم الرياضيـات ـ كـلية الرياضيات والحاسبات ـ جامعة الكوفة

الخـلاصـــة

دراسة صفات مصفوفات هيسيان ومصفوفات اليقين واستخدامها لايجاد النقاط الحرجة للاشكال التربيعية وتطبيقات الافتراضية الجديدة في مسائل الانظمة الديناميكية .

1- Introduction

In this paper, the calculation of critical points of quadratic forms are studied .The proposed method was applied in static springs system .

Hessian differential matrix was proved that it represents the symmetric matrix of the quadratic form.

Positive and negative matrices are play useful role in applied mathematical problems, Lyapunov stability[1], control system[2], electric circuit[3], economics[4], competition[5], approximation[6] and others.

In this paper the quadratic form

$$f(x) = c + b^t x + \frac{1}{2} x^t Ax \ , c \in R \ , b \in R^n , , x \in R^n , A \in R^{n \times n}$$

Was studied and proved that the stationary vector x^* is minimum only if $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ be positive definite ,and the gradient of quadratic form f(x) is $\nabla f(x) = b + Ax,$ and the Hessian matrix H of f(x) is the constant matrix A .

Then it is also proved that if A be positive definite proved that H is positive definite which implies that X* be minimum and calculated form

$$x^* = A^{-1} b$$

2- Basic Definitions and Theorems :

In this chapter, the basic definitions and theorems with properties are presented as follows :

Definition (2-1)[7] : the real symmetric n × n matrix $A=(a_{ij})$

is said to be **positive definite** if the leading principle minors of A are all positive , such that it satisfies the following Sylvestor's condition

$$\det(A_j)=\det(\begin{bmatrix} a_{11} & a_{12} & \dots & a_{ij} \\ a_{21} & a_{22} & \dots & a_{2j} \\ . & . & . & . \\ . & . & . & . \\ a_{j1} & a_{j2} & \dots & a_{jj} \end{bmatrix})>0 \quad , \quad \forall j=1,2,\dots,n$$

Definition (2-2) **[7]:** The real symmetric nxn matrix A is said to be **negative definite** if and only if

$(-1)^j \det(A_j)>0$ $\forall$ j=1,2, ...,n .

Theorem (2-1)[8-9] **:** Let $A=(a_{ij})$be n × n and symmetric, The following are equivalent.

1) $\det(A_j)>0$ $((-1)^j \det(A_j)>0),\forall j=1,2,\dots,n$

2) $x^t Ax>0$($x^t Ax<0$) for all vector x .

Then A is called positive (negative) definite .

Definition (2-3) **[10]:** A **quadratic from** on R is a real – valued function of the from $Q(x1,\dots\dots\dots xn)=\sum_{i\le j} a_{ij}x_i x_j$ where each term is a monomial of degree two , a matrix from Q(x)= x^t . A . x .

one can take the gradient of f(x) , as

$$\nabla f=\left[\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\dots\dots\dots,\frac{\partial f}{\partial x_n}\right]^t$$

Which is a column vector of n function .

Definition (2-4)[9,11] :-The n × n matrix of second derivatives is called the **Hessian matrix** and can be found by taking the gradient a gain , thus,

$$H(x_1,\dots,x_n)=\nabla\nabla^T f=\nabla^2 f$$

or more specifically, the ∇^2 operator can be written as

$$\nabla^2 = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} \\ . & & . \\ . & \frac{\partial^2}{\partial x_i \partial xj} & . \\ . & & . \\ \frac{\partial^2}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2}{\partial x_n^2} \end{bmatrix} = \left[\frac{\partial^2}{\partial x_i \partial x_j} \right]_{n\times n}$$

Where the subscript i denotes the row and j denotes the column.
The Hessian matrix is symmetric because of the properties of the second derivatives of continuous functions. $\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \frac{\partial^2}{\partial x_j \partial x_i} f(x)$

Definition (2-5)[10]:- A point x^{**} is said to be the **strong global minimum** of a function f(x) if $f(x) - f(x^{**}) > 0, \quad \forall x \in R^n$

Definition(2-6)**[10]** a point x^* is said to be a **strong local minimum** of a function f(x) if there exists a $\delta > 0$ such that
$f(x) - f(x^*) > 0$ for all x such that $\|x - x^*\| < \delta$.
If the inequalities in $f(x) - f(x^{**}) > 0$ and $f(x) - f(x^*) > 0$ are replaced by < than we have strong global and local maxima.
If the inequalities in $f(x) - f(x^{**}) > 0$ and $f(x) - f(x^*) > 0$ are replaced by $\geq$ then the minima are called weak global and weak local minima .

Definition(2-7)[10] One can expand f(x) in a **Taylor expansion** about x as

$f(x + \Delta x) = f(x) + \Delta x^T \Delta f(x) + \frac{1}{2} \Delta x^T \Delta^2 f(x) \Delta x + O(\|Ax\|^3)$ .Where the notation $O(\|Ax\|^3)$ represents terms " order " $\|Ax\|^3$ and these can be neglected for $\|Ax\|$ sufficiently small .

Theorem (2-1)[10,12] :- Necessary condition based on $(\nabla f, \nabla^2 f)$
x^* is a local minimum only if $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive semi definite .

Definition (2-8)[10,12]:- a stationary point which is not local minimum or local maximum is called a saddle point.

3-Proposed Method for Finding Critical Points : In this section the quadratic Form

$$f(x)=c+b^t x+\frac{1}{2}x^t Ax \ ,c\in R \ ,b\in R^n ,,x\in R^n ,A\in R^{n\times n} \qquad (3.1)$$

was studied and proved that the stationary vector $x*$ is minimum only if $\nabla f(x*)=0 \quad and \quad \nabla^2 f(x*)$be positive definite.

It is proved that the gradient of quadratic form $f(x)$ is $\nabla f(x)=b+Ax$ and the Hessian matrix H of f(x) is the constant matrix A.

Then if A be positive definite proved that H is positive definite which implies that x* be minimum and calculated from the form $X*=-A^{-1}b$.

Following a propositions are introduced :

Now a point x^* is called local minimum of $f(x)$ if there exist $\delta>0$ such that $f(x)-f(x^*)>0$ for all x such that $\|x-x^*\|<\delta$

where $\|.\|$ denotes the norm .

A point $x^*$is called global minimum of $f(x)$ if there exist $\delta>0$ such that $f(x)-f(x^*)>0$ for all x.

When the inequality in $f(x)-f(x^*)>0$ be replaced by $<$,then $x^*$is called maxima [11] .

One can expand $f(x)$ in a Taylor expansion about x as :

$$f(x+\Delta x)= f(x)+\Delta x^t\nabla f(x)+\frac{1}{2}\Delta x^t\nabla^2 f(x)\Delta x+O(\|\Delta x\|^3)$$

Where the notation $O\|\Delta x\|^3$ represents terms of order $\|\Delta x\|^3$ and these can be neglected for $\|\Delta x\|$ sufficiently small .

For x^* to be minimum this implies that $f(x^*+\Delta x)-f(x^*)>0$ for $\|\Delta x\|>0$.

From the Taylor expansion [11] :

$$f(x^*+\Delta x)= f(x^*)+\Delta x^t\nabla f(x^*)+\frac{1}{2}\Delta x^t\nabla^2 f(x^*)\Delta x$$

Must be greater than zero ,where $O\|\Delta x\|^3$ be neglected .

For x^*be stationary then

$\nabla f(x^*)=0$

Implying definition (2.1) that $\nabla^2 f(x^*)$ be positive definite .

Proposition (3.1) x^* is minimum only if $\nabla f(x^*)=0$ and $\nabla^2 f(x^*)$ be positive definite

Now for proposition(2) one can write f(x) in matrix notation as equation (3.1) >

In algebraic form one can write f(x) as

$$f(x) = c + \sum_{i=1}^{n} b_i x_i + \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{1}{2} x_j a_{ij} x_j$$

$$f(x) = c + \sum_{i=1}^{n} b_i x_i + \frac{1}{2}\left[a_{kk} x_k^2 + x_k \sum_{j \neq k} a_{kj} x_j + x_k \sum_{j \neq k} x_i a_{ik} . + \sum_{i \neq k} \sum_{j \neq k} x_i a_{ij} x_j \right]$$

And there fore

$$\frac{\partial f}{\partial x_k} = b_k + \frac{1}{2}\left[2a_{kk} x_k + 2\sum_{j \neq k} a_{kj} x_i \right] = b_k + \sum_{j} a_{kj} x_i$$

In matrix form

$$\nabla f(x) = b + Ax$$

Proposition(3.2):-

The gradient of quadratic form (3.1) is $\nabla f(x) = b + Ax$.

Now for new proposition, using definition (2.4), Hessian matrix be

$$H(x) = \nabla(\nabla f(x))' = \left[\frac{\partial^2}{\partial x_i \partial x_j} f(x) \right]_{n \times n}$$

Since $\frac{\partial f}{\partial x_j} = b_j + \sum_{k} a_{jk} x_k$

Then $\frac{\partial f}{\partial x_j x_i} f(x) = a_{ij} = a_{ji}$

And in matrix notation . H(x) = H = A

Proposition (3.3):-

The Hessian matrix of quadratic form is the constant matrix A. then.
If A be positive definite then H is positive definite which implies that x* be minimum.

Finally for introducing other proposition , the minimum x^* hold:

$$\nabla f(x^*) = b + Ax^* = 0 \Rightarrow x^* = -A^{-1} b$$

Proposition (3.4):-

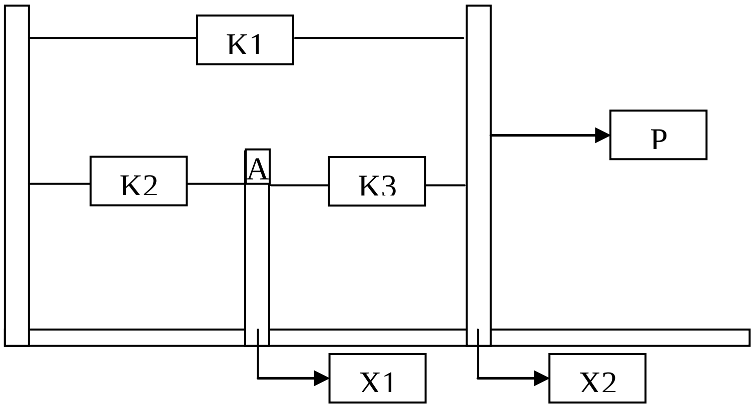
The minimum x* of quadratic form is calculated as

$$x^* = -A^{-1} b$$

4- **Applications of Hessian matrix**

This section is concerned for applications of Hessian matrix ,referred as linear spring system

To apply above new propositions, one can introduce the following System of linear spring system[13]:



The blocks A and B are two bodies connected to 3 Linear elastic springs having spring constants K_1,K_2,k_3 respectively .
when the force p= 0 ,$x_1 = 0$,$x_2 = o$ define the natural position .
One want to find new positions x_1 , x_2 when non zero P is applied [12 , 13 , 14].
Recall that for a linear spring blocks Law springs that $F = kd$ where k is spring constant and d is displacement of the spring from equilibrium.
The strain energy of a spring is the work done in stretching it,
that is

$$E = \int_0^x kxdx = \frac{1}{2}kx^2$$

The work put in to the system by the constant for P is given by the potential energy of the system is

$$f(x) = \frac{1}{2}k_1x_1^2 + \frac{1}{2}k_3(x_2 - x_1) + \frac{1}{2}x_2^2 - px_2$$

According to principle of potential energy f(x) must be minimize f(x) can be written in quadratic form :

$$f(x) = \frac{1}{2}[x_1 \; x_2]\begin{bmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [0 \; -p]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Using introduced propositions then

$$H = A = \begin{bmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{bmatrix}, x^* = -A^{-1}\begin{bmatrix} o \\ -p \end{bmatrix}$$

Using theorem (1) $\overline{A}=\begin{bmatrix}1 & \frac{-k_3}{k_2+k_3} \\ 0 & \frac{k_1+k_3-k_3^2}{k_2+k_3}\end{bmatrix}_{for\ k_1+k_3>k_3^2}$ is positive definite

And stationary $x^*=\frac{1}{|A|}\begin{bmatrix}-pk_3 \\ (k_2+k_3)p\end{bmatrix}$ be minimum

5- Conclusion

The general method for finding critical points of more one variable function is determined by partial derivatives .

In this method , the critical points studies by using definite and Hessian matrices , The Hessian matrix method involves multi variable Functions .

The advantage of this method is a obvious in its applications of ion , static springs systems .

Reffrences

1. Grujie,L.T.and siljack,D.D., "Asymptotic stability and Instability of Large – scale System" ; IEEE transaction automatic control, vol.Ac – 18 , pp.636.1973.
2. Ogatak, "modern control engineering" prentice ball 1998.
3. Shweehdi, M. *et al*, " Aparametric sensitive formulation for power system Analysis",14th pscc,sevilla,june,pp.24-28,2002.
4. MenzikI" Banach fixed point theorem and the stability of the market"International con ference inmathematics Brno, Czech, sept. pp.177-180,2003.
5. Burdern,I.r.f.fairas,J-D"Numerical Analysis"ITP C_o, 5e,1997.
6. Kolman B,1984,Introductory Linear Algebra with Applicatias Mac Millan publishing company, New York.
7. Barntt.S. and story,C." matrix methods in stability theory" Nelson, 1970.
8. Rizak A-"on positive and negative Definite matercies" J – Babylon un . vol.7,no.3,march,pp.472 – 475,2002.
9. Kaddora. I"Aprocedur for constru cting Lyaponur functions to sdtudy Dynamic systemso" Ms thesis .Bayhdadun.1984.
10. Brussel I J" calculus and Analytic Geometry"4e, Addison – Wesley pudlishing company; 1984.
11. Cinlar E & vanderby R – J ; mathemematical methods of Enyineering Company; 1984.
12. Kreyszige " Advanced Engineering Mathematical";M c Graw Hill;1984.
13. Meriamj I ; "Engineering mechanics static Dynamic"John wiley &sons, 1984.
