

Matrix Analysis with Applications
Dr. Sanjeev Kumar
Department of Mathematics
Indian Institute of Technology Roorkee

Lecture – 24
Gram Matrix and Minimization of Quadratic Forms

Hello friends. So, welcome to the 24th lecture of this course. So, in this lecture we will learn about gram matrix and we will learn how to find out a suitable minimizer for the given quadratic form.

(Refer Slide Time: 00:41)

Definition

Gram Matrix: Let A be a $m \times n$ matrix, then the $n \times n$ matrix

$$K = A^T A$$

is known as the associated Gram Matrix.

If $A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{bmatrix}$, then

$$K = A^T A = \begin{bmatrix} 6 & -3 \\ -3 & 45 \end{bmatrix}$$

is the Gram Matrix.

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 2

So, first definition in this lecture comes as a gram matrix. So, a gram matrix can be defined in this way let A be a m by n matrix. So, A can be a rectangular matrix also, then the n by n matrix a transpose into a equals to K is known as the associated gram matrix.

So from here, you can see since, it is a since. It is a product of the transpose of a matrix with that matrix itself. So, gram matrix will be always a symmetric matrix. If I take an example that, A is a 3 by 2 matrix. Where, first row of row of A is 1 3 2 0 and minus 1 6 then a transpose a comes out to be 6 minus 3 minus 3 and 45 .

So here a transpose, A will become 2 by 2 matrix and this is the gram matrix of A .

(Refer Slide Time: 01:59)

Theorem
All Gram matrices are positive semi-definite. The Gram matrix is positive definite $\iff \text{Ker}(A) = \{0\}$

Proof.
The associated quadratic form of K is $q(X) = X^T K X = X^T A^T A X = (AX)^T (AX) = \|AX\|^2 \geq 0 \forall x \in \mathbb{R}^n$. This implies K is positive semi-definite.
Moreover, it equals to 0 $\iff AX = 0$, and so if A has trivial kernel, this requires $X = 0$, and hence $q(X) = 0 \iff X = 0$. Thus in this case K is positive definite. \square

NT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 3

Now, we are having a very important result related to gram matrices. So, all gram matrices are positive semi definite, moreover the gram matrix is positive definite if and only if the kernel of A equals to 0, means the kernel is having only 0 vector.

So, let us see the proof of this.

(Refer Slide Time: 02:26)

Proof:
 $q(x) = X^T K X = X^T A^T A X = (AX)^T (AX) = \|AX\|_2^2 \geq 0$
 $\Rightarrow q(x) \geq 0 \Rightarrow K$ is positive semi definite.
2nd part \Rightarrow We need to prove that
 K is PD if and only if $\text{Ker}\{A\} = \{0\}$
Let $\text{Kernel}(A)$ contains only zero solⁿ
 $AX = 0$ only when $x = \{0\}$
 $q(x) \geq 0$ and $q(x) = 0$ only when $x = 0$
 $\Rightarrow K$ is PD
If K is PD \Rightarrow All the eigenvalues of K are positive.
 $\Rightarrow \text{rank}(A^T A) = n \Rightarrow \|AX\|_2^2 > 0$
 $\Rightarrow AX = 0$ only when $x = 0$

So, first we have to prove that all gram matrices are positive definite. So, if I take A So, if I take the quadratic form of a gram matrix K , it can be written as X transpose K into X . Where, K is the gram matrix. So, this can be defined as a transpose into A into X this can

be written as $A^T X$ transpose into a X . Which is the norm of which will be always greater than equals to 0.

So, since the quadratic form of a gram matrix will be always greater than equals to 0. So hence, K is positive semi definite; this is the first part of the theorem. Now, in the second part we need to prove that, the matrix K is positive definite if and only if kernel of the matrix A contains only 0 vector.

So, let kernel of A contains only 0 solution. In this case, the system AX equals to 0, implies that X equals to 0. Since, we are having that $q^T X$ greater than equals to 0 then and $q^T X$ equals to 0. Only when, X equals to 0 because, kernel is having only 0 vector. So, hence K is positive definite.

So, this we have done that, if kernel of A equals to 0, then K is positive definite. So, let us do the same result from the other way that is, if we assume that, if K is positive definite then we need to see. So, that the kernel of A is having only 0 vector. So, if K is positive definite means all the eigenvalues of K are positive. It means the rank of the matrix a transpose A is n . If K is $A^T A$ matrix means it is a full rank matrix.

Norm of a AX square is greater than equals to 0 or. In fact, it will be strictly greater than 0. So, hence AX will be 0 only when X equals to 0. This means the kernel of A contains only the 0 vector. So, this is the proof of other side.

(Refer Slide Time: 07:18)

The slide features a blue header with the word "Definition" in white. The main content area is white with a light blue background for the definition text. At the bottom, there is a blue footer containing logos for IIT Roorkee and NPTEL Online Certification Course, along with a small number "4".

Definition

Weighted Gram Matrix: If C is any symmetric, positive definite $m \times m$ matrix with all positive entries, then the weighted Gram Matrix is

$$K = A^T C A$$

IIT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 4

In the same way, we can define weighted gram matrix. So, if C is any symmetric positive definite m by m matrix with all positive entries. Then the weighted gram matrix is K equals to a transpose CA .

Weighted gram matrix is K equals to a transpose CA .

(Refer Slide Time: 07:45)

The slide is titled "Example" and contains the following text and equations:

Consider

$$C = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 5 \end{bmatrix}, \text{ and } A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 6 \end{bmatrix}$$

Then $K = A^T C A = \begin{bmatrix} 12 & -9 \\ -9 & 207 \end{bmatrix}$ which is positive definite.

Note: The theorem given on slide 3 also holds for weighted Gram matrix.

At the bottom of the slide, there are logos for "IIT ROORKEE" and "NPTEL ONLINE CERTIFICATION COURSE", and the number "5" in the bottom right corner.

So, for example, if I take C equals to this one, it is symmetric as well as positive definite matrix then and A is the same as we have taken in earlier case. Then, the associated gram matrix is A transpose $G C A$ which is given as 12 minus 9 minus 9 and 2 0 7 which is again positive definite.

So, all the result which we have discuss in the earlier theorem holds for weighted gram matrix also. Now, we will look for the minimization of quadratic forms.

(Refer Slide Time: 08:26)

Minimization of Quadratic forms

Consider a scalar quadratic function $p(x) = ax^2 + 2bx + c \quad \forall x \in \mathbb{R}$

1 If $a > 0$, then the graph of p is a parabola pointing upwards, and so \exists a unique minimum value, i.e by elementary calculus


$$\frac{\partial p}{\partial x} = 2ax + 2b = 0 \implies x^* = -\frac{b}{a} \text{ as a critical point}$$

Now $\frac{\partial^2 p}{\partial x^2} = 2a > 0$ as a is positive.

$$\implies x^* = -\frac{b}{a} \text{ is a point of minima and the minimum value is } p(x^*) = \frac{ac - b^2}{a}$$

2 If $a < 0$, then $\frac{\partial^2 p}{\partial x^2} = 2a < 0$

$$\implies x^* = -\frac{b}{a} \text{ is a point of maxima and the maximum value is } p(x^*) = \frac{ac - b^2}{a} \text{ which is validated since } a < 0, p \text{ is a parabola pointing downwards, and so } \exists \text{ a unique global maximum.}$$

 NPTEL ONLINE CERTIFICATION COURSE 6

So, consider a scalar quadratic form px equals to ax^2 plus $2bx$ plus c for all x belongs to \mathbb{R} . So, here what we are having we are having a .

(Refer Slide Time: 08:41)

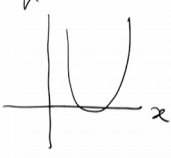
$$q(x) = ax^2 + 2bx + c$$

① $a > 0 \implies q'(x) = 2ax + 2b = 0 \implies x = -\frac{b}{a}$ critical point $q(x)$

$$q''(x) = 2a > 0$$

So point $x = -\frac{b}{a}$ is a point of minima.

② $a < 0 \implies x = -\frac{b}{a}$ is point of maxima.



Scalar quadratic form qx equals to ax^2 plus $2bx$ plus c . Now, consider the case, when a is greater than 0. So, when a is greater than 0 it will be a parabola pointing upward and if I see the minimization of these.

So, it will be having an unique global minima. So, if I see this so, q' dash x will become $2x$ plus $2b$ equals to 0 which gives me 0 equal x equals to minus b upon a is the critical

point. Now, if I check the sufficient condition of minimization on this, then the double derivative of q with respect to x can be given as 2 twice of a and if a is positive then twice of a is also positive. So, point x equals to $-\frac{b}{a}$ is a point of minima. Similarly,, if I take a is less than $- 0$. In this case, the point x equals to $-\frac{b}{a}$ comes out to be a point of maxima.

So, generally these are the necessary and sufficient conditions we can use for looking the maxima and minima of a scalar quadratic form, but if I am having a vector quadratic form, where x is not a scalar, but it is a vector form n . Dimensionally space then, how to find out the minima of the quadratic form? So, we will learn this application now.

(Refer Slide Time: 11:18)

General Case

Consider a general quadratic function of n variables x_1, x_2, \dots, x_n as



$$p(X) = p(x_1, x_2, \dots, x_n) = \sum_{i,j=1}^n k_{ij}x_i x_j - 2 \sum_{i=1}^n f_i x_i + c$$

the coefficients k_{ij}, f_i, c are assumed to be real.

In compact form

$$p(X) = X^T K X - 2X^T f + c$$

where $K = \{k_{ij}\}_{n \times n}$ is a symmetric matrix, f is a constant vector, and c is a constant scalar.



7

So, consider a general quadratic function of n variables x_1, x_2, \dots, x_n as $p(X)$ where X is n dimensional vector having component x_1, x_2, \dots, x_n and this can be written as $\sum_{i,j=1}^n k_{ij} x_i x_j - 2 \sum_{i=1}^n f_i x_i + c$.

Where, the coefficient k_{ij}, f_i, c are assumed to be real. So in compact form, this particular general quadratic function can be written in by the matrix notations. So, here it can be written as $X^T K X - 2X^T f + c$.

(Refer Slide Time: 12:25)

$$\begin{aligned}
p(x_1, x_2, x_3) &= 2x_1^2 + 3x_2^2 - 5x_3^2 - 2x_1x_2 + 2x_2 + 3 \\
&= [x_1 \ x_2 \ x_3] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - 2[x_1 \ x_2 \ x_3] \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + 3 \\
K &= \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & -5 \end{bmatrix} \text{ and } f = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}; \quad c = 3 \\
p(x) &= x^T K x - 2x^T f + c
\end{aligned}$$

Where K is a symmetric matrix f is a constant vector and C is a constant scalar. For example, if you are having a general quadratic form of 3 variables that is let us say 2x1 square plus 3 x 2 square minus 5 x 3 square minus 2x 1 x 2 plus 2 times x 2 plus 3.

Then in matrix form I can write it as x1 x 2 x 3 multiplied with the symmetric matrix 2 3 minus 5 and then minus 1 minus 1 because the coefficient of x1 x2 is minus 2 the coefficient of x 2 x 3 as well as x one x 3 are 0. So, all these entries will be 0 into x1 x2 x3.

Then what I am having minus 2 times minus 2x T f. So, x T will become again x 1 x 2 x 3 and f will be here a column vector 0 minus 1 0 plus 3. So, if I see here, K comes out to be 2 minus 1 0 minus 1 3 0 and 0 0 minus 5 and the vector f is 0 minus 1 0 the scalar C is 3. So, this form p X can be written as XTKX minus 2 times XT f plus C. So, in this way we can write any quadratic function of n variables into a compact form by using this concept.

(Refer Slide Time: 14:50)

Theorem

If K is a symmetric Positive Definite Matrix, then the quadratic function

$$p(X) = X^T K X - 2X^T f + C$$

has a unique minimizer, which is the solution to the linear system

$$KX = f \text{ namely } X = K^{-1}f$$

The minimum value of $p(X)$ is equal to

$$p(X^*) = p(K^{-1}f) = C - f^T K^{-1}f = C - f^T X^* = C - (X^*)^T K X^*$$

IIT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 8

Now, this is the theorem which tells us about the minimization of a general quadratic form. So, if K is a symmetric positive definite matrix then the quadratic function $p(X)$ equals to X transpose KX minus twice of $X^T f$ plus C has a unique minimize.

So, please note down if the unique minimization or global minima of this particular quadratic function depends on the property of K . So, if K is symmetric and positive definite then this particular quadratic function will certainly have a global minima.

And this global minima is the solution of the linear system KX equals to f namely as X equals to K inverse f . The minimum value of $p(X)$ is equals to in any one of this form. So, $p(X^*)$ equals to $p(K^{-1}f)$ which is C minus f transpose K inverse f . So, this K inverse f I can replace with the X^* which is the K solution of this system. So, C minus f transpose X^* or this also I can write C minus $X^* T K X^*$. So, let us see the proof of this.

(Refer Slide Time: 16:22)

Proof: Since K is a PD matrix $\Rightarrow K$ is non-singular and hence $X^* = K^{-1}f$ exists.

Now for any $X \in \mathbb{R}^n$, we can write

$$\begin{aligned} p(X) &= X^T K X - 2X^T f + C \\ &= X^T K X - 2X^T K X^* + C \\ &= (X - X^*)^T K (X - X^*) + [C - (X^*)^T K X^*] \\ X^T K X^* &= (X^*)^T K X \end{aligned}$$

\rightarrow

So, since K is a positive definite matrix. This implies K is nonsingular and hence X^* equals to K inverse f exists, it means K is invertible. Now, for any X belongs to n dimensional real vector space. We can write p of X equals to X transpose KX minus twice $X^T f$ plus C . This equals to $X^T K X$ minus $2X^T K X^*$ and from here, I can write f equals to K into X^* plus C .

Or this particular expression, I can write also X minus X^* transpose KX minus X^* plus C minus X^* transpose KX^* only writing this I have use the symmetry of K because K is a symmetric matrix. That is K equals to K transpose to identify X transpose KX^* equals to X^* transpose K into X . So, here I have use the symmetry of K .

(Refer Slide Time: 18:58)

Proof: Since K is a PD matrix $\Rightarrow K$ is non-singular and hence $X^* = K^{-1}f$ exists.

Now for any $X \in \mathbb{R}^n$, we can write

$$\begin{aligned} p(X) &= X^T K X - 2X^T f + C \\ &= X^T K X - 2X^T K X^* + C \\ &= \underline{(X - X^*)^T K (X - X^*)} + \underline{[C - (X^*)^T K X^*]} \end{aligned}$$

The first term in the above expression has the form $Y^T K Y$, where $Y = X - X^*$

Since K is PD, we know $Y^T K Y > 0$ for all $Y \neq 0$. Thus, the first term achieves its minimum value to zero if and only if $Y = 0 \Rightarrow X - X^* = 0 \Rightarrow X = X^*$. Moreover, since X^* is fixed, the second term does not depend on X . Therefore, the minimum value of $p(X)$ occurs at $X = X^*$. The minimum value of $p(X)$ can be achieved by putting $X = K^{-1}f$ in $p(X)$. //

The first term means this term. In the above expression, has the form $Y^T K Y$ where Y equals to X minus X^* . Since, K is positive definite. We know $Y^T K Y$ will be greater than will be a strictly greater than 0. For all Y those are not equals to 0. Thus, the first term means the underlined term achieves. It is minimum value because, it is a quadratic term to 0 if and only if, Y equals to 0 because, then only this term will become 0 and this implies Y equals to X minus X^* . So, X minus X^* equals to 0 or X equals to X^* .

Moreover, since X^* is fixed the second term does not depend on X because X^* is fixed here and this term does not contain any expression in X . Therefore, the minimum value of $p(X)$ occurs at X equals to X^* ok. And where X^* equals to $K^{-1}f$ which is the proof of first part.

And the minimum value is or can be obtained just by substituting X equals to $K^{-1}f$ in this expression and from that we can get any form. So, the minimum value of $p(X)$ can be achieved by putting X equals to $K^{-1}f$ in $p(X)$ means, in this expression, which will give you the second part of the second statement of the theorem proof of the second statement of the theorem.

So, this is the proof of this theorem. Now, let us see an example of this theorem. So, example is like this.

(Refer Slide Time: 23:29)

Ex: Consider the problem of minimizing the quadratic function
 $p(x_1, x_2) = 4x_1^2 - 2x_1x_2 + 3x_2^2 + 3x_1 - 2x_2 + 1$
 over \mathbb{R}^2 .

Solⁿ: By using the previous theorem,

$$p = [x_1 \ x_2] \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 2[x_1 \ x_2] \begin{bmatrix} -\frac{3}{2} \\ 1 \end{bmatrix} + 1$$

$$* p = x^T K x - 2x^T f + c$$

$$K = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \Rightarrow K \text{ is PD} \quad x^* = K^{-1}f$$

$$f = \begin{bmatrix} -3/2 \\ 1 \end{bmatrix} \quad = \begin{pmatrix} -0.31818 \\ 0.22727 \end{pmatrix}$$

$$c = 1 \quad p(x^*) = \frac{13}{44}$$

Consider the problem of minimizing the quadratic expression or quadratic function $p(x_1, x_2)$ equals to $4x_1^2 - 2x_1x_2 + 3x_2^2 + 3x_1 - 2x_2 + 1$ over \mathbb{R}^2 means all real values of x_1 and x_2 .

So, by using the. So, how to solve this particular thing? So, by using the previous theorem which we have stated just now if the associated matrix K will be positive and symmetric positive definite and symmetric. Then, it will be having a unique minima. So, let us write this p in the matrix form. So, it will become $x_1 \ x_2$ and then the matrix 4 minus 1 minus 1 and 3 and then $x_1 \ x_2$ then I am having minus 2 times x^T into f . So, x^T is $x_1 \ x_2$ and what will be f here. So, if we see the coefficient of x_1 it is plus 3 . So, it will become minus 3 by 2 and if we see the coefficient of x_2 it is minus 2 . So, it comes out to be 1 plus 1 .

So, here what I am having, if I compare with $X^T K X - 2X^T f + C$. So, K is 4 minus 1 minus 1 and 3 f is minus 3 by 2 and 1 and C is 1 . If we see the matrix K , first principal minor of K is positive and the determinant K is 13 which is again positive. So, here K is positive definite. So, since K is positive definite, so according to the previous theorem the quadratic function p will be having a unique minima.

Now, this particular function will be having a unique minima at X^* X^* equals to $K^{-1}f$. So now, if I calculate the K inverse, by the previous theorem X^* equals to $K^{-1}f$ and this comes out to be minus 0.31818 and 0.22727 .

So, this is the point on which the minima of this function will exist and the minimum value of this comes out to can be calculated using Any of the expression of this theorem means, either $C - f^T K^{-1} f$ or $C - f^T X^*$ or $C - X^{*T} K X^*$.

(Refer Slide Time: 28:04)

Example

Minimize the following quadratic function
 $p(x_1, x_2) = 4x_1^2 - 2x_1x_2 + 3x_2^2 + 3x_1 - 2x_2 + 1$ over all real values of x_1 and x_2 .

$$p(x_1, x_2) = [x_1 \ x_2] \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 2 [x_1 \ x_2] \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} + 1, \text{ Here}$$

$$K = \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}; f = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$$

Hence $X^* = K^{-1}f = \begin{bmatrix} -7 \\ 22 \\ 22 \end{bmatrix}$

Hence, $p(X^*) = 1 - \begin{bmatrix} -7 & 5 \\ 22 & 22 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -7 \\ 22 \\ 22 \end{bmatrix} = \frac{13}{44}$

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 10

(Refer Slide Time: 28:04)

Proof

Proof.

Since K is PD $\implies K$ is non-singular. Hence $X^* = K^{-1}f$ is well defined .
 Moreover for any $X \in \mathbb{R}^n$, we have

$$p(X) = X^T K X - 2X^T f + C = X^T K X - 2X^T K X^* + C = (X - X^*)^T K (X - X^*) + [C - (X^*)^T K X^*]$$

Minimum value of first term is zero $\iff X - X^* = 0 \implies X = X^*$ is point of minima. \square

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 9

(Refer Slide Time: 28:05)

Theorem
If K is a symmetric Positive Definite Matrix, then the quadratic function

$$p(X) = X^T K X - 2X^T f + C$$

has a unique minimizer, which is the solution to the linear system

$$KX = f \text{ namely } X = K^{-1}f$$

The minimum value of $p(X)$ is equal to

$$p(X^*) = p(K^{-1}f) = C - f^T K^{-1}f = C - f^T X^* = C - (X^*)^T K X^*$$

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 8

Where, X^* is given by this vector and this minimum value is 13 by 44. So, this is the point of minima and this is the minimum value of p at this point of minima. So, in this way by using the previous theorem, we can find out the unique minimizer of a given quadratic function; however, the condition as I stated earlier the matrix K should be positive definite.

(Refer Slide Time: 28:58)

References

- 1 Hoffman, K. and Kunze, R., Linear Algebra, second edition, Pearson Education (Asia) Pvt. Ltd /Prentice Hall of India, 2004
- 2 Leon, S.J., Linear Algebra with Applications, 8th Edition, Pearson, 2009
- 3 Strang, G., Linear Algebra and its Applications, 3rd edition, Thomson Learning Asia Pvt Ltd, 2003
- 4 Meyer C. D., Matrix Analysis and Applied Linear Algebra, ISBN-10: 0898714540

IT ROORKEE NPTEL ONLINE CERTIFICATION COURSE 13

So, with these in this lecture we have learnt gram matrix and how to minimize a given quadratic function of n variables.

In the next lecture, we will learn another canonical transformation that is called Jordan canonical form. Which exist when a matrix is not diagonalizable? So, with this I will end this particular lecture. These are the references for this lecture.