# Proof of a matrix is positive semidefinite iff it can be written in the form $X'X$

I know the fact that a matrix is positive semidefinite if and only if it can be written in the form $X'X$. But how to prove it?

Thanks in advance.

(matrices)

asked Sep 3 '13 at 3:22
Ian
579 3 15

1 It would be useful to know how you came across the problem and whether or not you have tried to solve it yourself, and if so what you did and why/where you still need help. – Ben Millwood Sep 3 '13 at 3:30

1 @Ian: I take it by $X'$ you mean what is usually written $X^T$, the adjoint or transpose of $X$, i.e., $X^T_{ij} = X_{ji}$, where the $X = [X_{ij}]$? – Robert Lewis Sep 3 '13 at 4:28

Here you go, straight from Wikipedia. – Omnomnomnom Sep 3 '13 at 4:35

Use the spectral theorem to reduce to the diagonal case. – Qiaochu Yuan Sep 3 '13 at 7:56

## 1 Answer

OK, first off, I am assuming the notation $X'$ our OP Ian uses in his question is in fact, as I asked in my comment, the same as the conventional $X^T$, which I shall deploy throughout.

It is easy to see that a matrix $A$ such $A = X^T X$ is positive definite; we have

$$\langle v, Av\rangle = \langle v, X^T X v\rangle = \langle Xv, Xv\rangle \geq 0. \tag{1}$$

However, going the other way is much more difficult. And the reason it is much more difficult is:

**IT IS FALSE** that every positive semi-definite matrix $A$ can be written $A = X^T X$.

However,

**IT IS TRUE** that every *symmetric* positive semi-definite matrix $A$ can be so written.

To see this, suppose $A = A^T$; then $A$ may be diagonalized by some orthogonal matrix $O$: $O^T A O = \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, where $n$ is the size of $A$. Then for any vector $x$, set $y = O^T x$, so that $Oy = OO^T x = x$; then

$$\langle x, Ax\rangle = \langle Oy, AOy\rangle = \langle y, O^T AOy\rangle = \langle y, \Lambda y\rangle, \tag{2}$$

so that taking $y = (y_1, y_2, \dots, y_n)^T$ yields

$$\langle x, Ax\rangle = \sum_1^n \lambda_i y_i^2; \tag{3}$$

this shows we must have all $\lambda_i \geq 0$, hence we can take

$$\Lambda^{\frac{1}{2}} = \operatorname{diag}\left(\lambda_1^{\frac{1}{2}}, \lambda_2^{\frac{1}{2}}, \dots, \lambda_n^{\frac{1}{2}}\right), \tag{4}$$

and setting $X = O\Lambda^{\frac{1}{2}} O^T$, we see that

$$X^T X = O^{TT}\left(\Lambda^{\frac{1}{2}T} O^T O \Lambda^{\frac{1}{2}}\right) O^T = O\Lambda O^T = A, \tag{5}$$

establishing the desired result. Note we have exploited the symmetry of $\Lambda^{\frac{1}{2}}$ in performing this calculation.

It should be observed that $A = X^T X$ implies $A$ symmetric: $A^T = (X^T X)^T = X^T X^{TT} = X^T X = A$; thus a positive semi-definite $A$ would automatically be symmetric if the assertion $A = X^T X$ were true.

To see that $A$ positive semi-definite is not sufficient to force $A = X^T X$, consider the $2 \times 2$ matrix

$$A = \begin{bmatrix} \lambda & \epsilon \\ 0 & \lambda \end{bmatrix}, \tag{6}$$

and let $v = (x, y)^T$. Then

$$\langle v, Av \rangle = (x, y) \begin{bmatrix} \lambda & \epsilon \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda(x^2 + y^2) + \epsilon xy \tag{7}$$

I claim that, at least for $\lambda > 0$ and $|\epsilon| \le \lambda$,

$$\lambda(x^2 + y^2) + \epsilon xy \ge 0 \tag{8}$$

for all $(x, y)^T$. To see this, note that we can, without loss of generality, take $|x| \le |y|$, since the quadratic form $\lambda(x^2 + y^2) + \epsilon xy$ is symmetric in $x$ and $y$. Then we have

$$|\epsilon xy| = |\epsilon||x||y| \le \lambda |y|^2 = \lambda y^2, \tag{9}$$

or

$$-\lambda y^2 \le \epsilon xy \le \lambda y^2, \tag{10}$$

whence

$$0 \le \lambda y^2 + \epsilon xy \le \lambda(x^2 + y^2) + \epsilon xy, \tag{11}$$

showing that $A$ is positive semi-definite. But there is no similarity transformation which diagonalizes $A$; indeed, if for some nonsingular $W$, $WAW^{-1}$ *were* diagonal, we would have to have $WAW^{-1} = \lambda I$, since the characteristic polynomial of a matrix is a similarity invariant. But then we would have $W(A - \lambda I)W^{-1} = 0$; but this is impossible since $A - \lambda I = \epsilon N$, where $N = [\delta_{i,i+1}]$ is a non-vanishing, albeit nilpotent, matrix; it cannot be "similaritized" into vanishing.

In closing: certain generaizations may be possible: if $A$ is an $n \times n$ matrix of the form

$$A = \lambda I + \epsilon N, \tag{12}$$

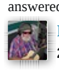
then $A$ is positive semi-definite provided $|\epsilon| \le \frac{\lambda}{n-1}$ for $\lambda > 0$; furthermore $A$ can't be diagonalized; the arguments are essentially the same as those given above.

Sorry about all the false starts, false alarms, and repetitive editing.

Hope this helps. Cheers.

edited Sep 14 '13 at 18:01

answered Sep 3 '13 at 10:26



Robert Lewis
**21.8k** 2 14 36

1 Blah! hit that "Post" button by mistake again! Damn this 'droid! Damn these beefy, guitar pickin' fingers! But hey, 's'all good! I'll just edit this thing into submission! Stay tuned . . . ;) – Robert Lewis Sep 3 '13 at 10:31

Blast! Did it again! Almost there, hang on, final edits pending! – Robert Lewis Sep 3 '13 at 17:43

I tried fixing the LaTeX, please check to ensure I did not add an error. – Avraham Sep 3 '13 at 17:51

@Av'racham: wil do. – Robert Lewis Sep 3 '13 at 19:45

Still needs a little more work. Comong, coming . . . SOON! – Robert Lewis Sep 3 '13 at 19:53