# Weakly positive definite matrices

```
class point (x,y); real x,y;
   begin ref (point) procedure plus (P); ref (point) P;
               plus :- new point (x+P.x, y+P.y);
   end point;

point class polar;
   begin real r,v;
         ref (polar) procedure plus (P); ref (point) P;
               plus :- new polar (x+P.x, y+P.y);
         r := sqrt (x↑2+y↑2);
         v := arctg (x,y)
  end polar;
```

Trygve Kastberg Nilssen

[ simula . research laboratory ]

Research Report Simula 2005-07

Visiting address:
Martin Linges vei 17, Fornebu

P.O. Box 134
NO-1325 Lysaker, Norway

Telephone: +47 67 82 82 00
Telefax: +47 67 82 82 01

www.simula.no

# WEAKLY POSITIVE DEFINITE MATRICES

TRYGVE KASTBERG NILSSEN

Abstract. A weakly positive definite matrix is defined to be a matrix which can be written as a product of two positive definite matrices. This paper proves that a matrix is weakly positive definite if and only if the real eigenvalues are positive. The main components of the proof are Schur decomposition and mathematical induction.

## 1. Introduction

This paper proves the following theorem

**Theorem 1.1.** *Let $A$ be a nonsingular real matrix. The following two statements are equivalent*

$$\lambda(A) \cap \mathbb{R}^- = \emptyset \tag{1}$$

$$\exists\, B, C \text{ positive definite, such that } A = BC. \tag{2}$$

Here $\lambda(A)$ denotes the spectrum of $A$, which is the set of its eigenvalues. $\mathbb{R}^-$ is the set of strictly negative real numbers.

Positive definite matrices are defined by

**Definition 1.1.** *A real matrix $B$ is called positive definite if*

$$x^T Bx > 0, \tag{3}$$

*for all real and nonzero vectors $x$.*

Notice that $B$ is not assumed to be symmetric. We refer to $x^T Bx$ as the quadratic form of $B$.

We now introduce the term weakly positive definite as

**Definition 1.2.** *A matrix is called weakly positive definite if it can be factorized into two positive definite matrices.*

Obviously a positive definite matrix is also weakly positive definite, because it can be factorized into the identity and itself. The class of positive definite matrices is not closed under multiplication, so the class of weakly positive definite matrices is strictly larger.

Theorem 1.1 states that the weakly positive definite matrices are precisely characterized as the real and invertible matrices where the arguments of the eigenvalues differ from $\pi$.

The interest in weakly positive definite matrices arose through the study of higher–order discretization methods for time dependent PDEs, see [3, 4]. The PDEs are discretized in time by fully implicit Runge–Kutta methods, and the linear systems to be solved are preconditioned with block diagonal preconditioners. If the Runge–Kutta coefficient matrices were positve definite, which they usually are not in the fully implicit case, the preconditioned systems can be proved to have bounded condition numbers. However, as shown in [3, 4], it is sufficient that the Runge–Kutta coefficient matrices can be factorized into two positive definite matrices. Now, since A–stable Runge–Kutta matrices can easily be shown to not have negative eigenvalues, Theorem 1.1 is valuable.

This paper is organized as follows: Section 2 makes some remarks about weakly positive definite matrices. Section 3 proves Theorem 1.1.

## 2. Remarks

In the literature positive definite matrices are frequently assumed to be symmetric in addition to (3). It can be shown that symmetric positive definite matrices are characterized as symmetric matrices having positive eigenvalues (see [1, 2]).

In order to characterize nonsymmetric positive definite matrices we define the symmetric part of a matrix $B$ as

$$S(B) = \frac{1}{2}(B + B^T).$$

Since $x^T Bx$ is a scalar, we see that $x^T Bx = (x^T Bx)^T = x^T B^T x$, which means that transposing a matrix does not change its quadratic form. Thus we see that

$$x^T Bx = x^T S(B)x.$$

Therefore we can characterize (possibly nonsymmetric) positive definite matrices as matrices where the symmetric part has positive eigenvalues. By Theorem 1.1 weakly positive definite matrices are also characterized by their eigenvalues.

Symmetric matrices are known to have real eigenvalues, see e.g. [1, 2]. Thus

**Remark 2.1.** *In the symmetric case weakly positive definite and positive definite matrices are characterized equally.*

We now make two remarks which further motivates the word weakly in Definition 1.2.

**Remark 2.2.** *Weakly positive definite matrices are invariant under arbitrary change of basis, i.e. for an arbitrary invertible $P$, $PAP^{-1}$ is weakly positive definite if and only if $A$ is weakly positive definite.*

*Proof.* The result follows by Theorem 1.1 since $A$ and $PAP^{-1}$ are similar matrices and share the same set of eigenvalues, cf. e.g. [1, 2]. □

**Remark 2.3.** *Positive definite matrices are invariant under orthogonal change of basis, but not under arbitrary change of basis.*

*Proof.* Let $Q$ be an orthogonal matrix representing an orthogonal change of basis, i.e. $Q^T = Q^{-1}$. If

$$x^T Bx > 0, \quad \forall x \neq 0,$$

then by defining $y = Q^T x$ we see that

$$x^T QBQ^T x = y^T By > 0, \quad \forall y \neq 0.$$

To show that positive definite matrices are not invariant under an arbitrary change of basis, consider the standard $2 \times 2$ rotation matrix with an angle $\theta$, given by

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{4}$$

Then $R_\theta$ is positive definite if $\theta \in (-\pi/2, \pi/2)$. This can be seen because

$$x^T R_\theta x = cos(\angle(x, R_\theta x))\|x\|\|R_\theta x\| = cos\theta\, \|x\|^2.$$

Now, let $P$ be the positive definite matrix $P = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. Then $PR_{\pi/3}P^{-1}$ is easily seen to not being positive definite by choosing $x = (1,1)^T$

$$x^T PR_{\pi/3}P^{-1}x = 2\cos\frac{\pi}{3} - \frac{3}{2}\sin\frac{\pi}{3} < 0.$$

□

We end this section by giving an equivalent characterization of weakly positive definite matrices, which could have served as an alternative definition to Definition 1.2. The new characterization is closer to how Theorem 1.1 is used in practice, cf. [3, 4]. It will also be useful in the proof of Theorem 1.1.

**Lemma 2.1.** *A real matrix $A$ is weakly positive definite if and only if there exists a positive definite matrix $B$ such that the product $BA$ is positive definite.*

*Proof.* We start by observing that a matrix is positive definite if and only if its inverse is positive definite[1]. This is seen by the change of basis $x = B^{-1}y$:

$$x^T Bx = y^T B^{-T} BB^{-1}y = y^T B^{-T} y = y^T B^{-1}y.$$

If there exists a $B$ as in Lemma 2.1, we then see that $A$ can be factorized into the factors $B^{-1}$ and $BA$ which both are positive definite, and then $A$ is weakly positive definite. Further, if $A$ is weakly positive definite, let $B$ be as defined in Definition 1.2. Then both $B^{-1}$ and $B^{-1}A$ $(= C)$ are positive definite, and the lemma is proved.

□

## 3. Proof of Theorem 1.1

In this section we prove Theorem 1.1. First we need two lemmas. The first states that $2 \times 2$ matrices with truly complex eigenvalues are weakly positive definite.

**Lemma 3.1.** *Any $A \in \mathbb{R}^{2\times 2}$ with nonreal eigenvalues is weakly positive definite.*

*Proof.* In this proof we construct a positive definite $B$, such that $BA$ is positive definite. To do so, we study the angle between $x$ and $Ax$, written

$$\alpha(x) \equiv \angle(x, Ax) : \mathbb{R}^2 \to (-\pi, \pi].$$

[1]In other words, the property of being positive definite is closed under inversion.

Notice that if there exists an $x$ such that $\alpha(x) = 0$, then $A$ will have a positive real eigenvalue, which is against the assumption. Similarly, if there exists an $x$ with $\alpha(x) = \pi$, then there will be a negative eigenvalue of $A$. Thus $\alpha(x) \notin \{0, \pi\}$ for $x \neq 0$.

Notice further that $\alpha(x)$ is a continuous function of $x \neq 0$. This gives that for a given $A$ we will either have

$$\alpha(x) \in (0, \pi) \quad \text{or} \quad \alpha(x) \in (-\pi, 0) \quad \forall x \neq 0.$$

Assume that

$$\alpha(x) \in (0, \pi), \quad \forall x \neq 0,$$

and let

$$\theta = \sup_{x \neq 0} \alpha(x).$$

In the following we show that $\theta < \pi$. We have that $\alpha(x) = \angle(x, Ax) = \angle(\frac{x}{\|x\|}, \frac{Ax}{\|x\|})$, because changing the length of vectors does not change the angle between them. Therefore

$$\sup_{x \neq 0} \alpha(x) = \sup_{\|x\|=1} \alpha(x).$$

Since the set defined by $\|x\| = 1$ is a compact set, and $\alpha(x)$ is a continuous function of $x$, we can conclude that the supremum value is attained and that it is less than $\pi$, i.e.

$$\theta = \max_{x \neq 0} \alpha(x) \in (0, \pi).$$

Now, set $B = R_{-\theta/2}$ where $R_{-\theta/2}$ is the rotation matrix with an angle $-\frac{\theta}{2}$, c.f. (4). Since $B$ is a pure rotation matrix, we have that

$$\angle(x, Ax) \in (0, \theta), \quad \forall x \neq 0,$$
$$\Downarrow$$
$$\angle(x, BAx) \in (-\frac{\theta}{2}, \frac{\theta}{2}) \subset (-\frac{\pi}{2}, \frac{\pi}{2}), \quad \forall x \neq 0.$$

And further since $\angle(x, Bx) = -\frac{\theta}{2} \in (-\frac{\pi}{2}, 0)$, $\forall x \neq 0$, we get

$$x^T Bx > 0, \quad \forall x \neq 0,$$
$$x^T BAx > 0, \quad \forall x \neq 0,$$

which means that both $B$ and $BA$ are positive definite.

Finally, if

$$\alpha(x) \in (-\pi, 0), \quad \forall x \neq 0,$$

we define $\theta = \inf \alpha(x)$ and the result can be proved similarly. □

The next lemma proves that a special block matrix can be made positive definite. This will be used as an induction step for extending the above $2 \times 2$–result to Theorem 1.1.

**Lemma 3.2.** *Let $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$ be positive definite. For any $C \in \mathbb{R}^{n \times m}$ there exists an $\epsilon > 0$ such that the block matrix*

$$E = \begin{pmatrix} \epsilon A & \epsilon C \\ 0 & B \end{pmatrix}$$

*is positive definite.*

*Proof.* Since $S(A)$ is symmetric positive definite, $S(A)^\gamma$ is defined for $\gamma \in \mathbb{R}$, cf. [1, 2]. Further let

$$\alpha = \min_{\|x\|=1} x^T A x,$$
$$\beta = \min_{\|y\|=1} y^T B y.$$

Both $\alpha$ and $\beta$ exist and are positive because the objective functions are continuous and positive plus that the admissible sets are compact.

Let $z^T = (x^T, y^T)$. We have

$$\begin{aligned}
z^T E z &= \epsilon x^T S(A) x + \epsilon x^T C y + y^T B y \\
&= \|(\epsilon S(A))^{\frac{1}{2}} x + \frac{\epsilon}{2}(\epsilon S(A))^{-\frac{1}{2}} C y\|^2 - \frac{\epsilon}{4} y^T C^T (S(A))^{-1} C y + y^T B y \\
&\geq -\frac{\epsilon}{4} y^T C^T A^{-1} C y + y^T B y \\
&\geq \left(\beta - \frac{\epsilon}{4}\frac{\|C\|^2}{\alpha}\right)\|y\|^2 \qquad (5)\\
&> 0,
\end{aligned}$$

when $\epsilon < \frac{4\alpha\beta}{\|C\|^2}$ and $y \neq 0$. Here (5) follows by the definitions of $\alpha$ and $\beta$. $\square$

We are now ready to prove Theorem 1.1. In Lemma 3.2 $m$ and $n$ are arbitrary positive integers, but in the following proof only $n \leq 2$ is needed.

*Proof of Theorem 1.1.* First, we prove that a weakly positive definite matrix can not have negative eigenvalues, i.e. (2) $\Rightarrow$ (1). To see this, assume that $B$ is positive definite and that $A$ has a negative eigenvalue, $Ax = -\lambda x$, $\lambda \in \mathbb{R}^+$ where $x$ is an eigenvector. Then

$$x^T B A x = -\lambda x^T B x < 0,$$

and therefore $BA$ is not positive definite.

Next, we show that if $A$ has no real negative eigenvalues, then $A$ is weakly positive definite, i.e. (1) $\Rightarrow$ (2). This is done by constructing a positive definite matrix $B$, such that $BA$ is positive definite. The construction utilizes Schur decomposition.

Let $A = QTQ^T$ be the real Schur decomposition of $A$, where $Q$ is orthogonal and $T$ is a block[2] upper triangular matrix with $1 \times 1$ or $2 \times 2$ blocks on the diagonal, see [1]. The real eigenvalues of $A$, which are positive, can be found on the $1 \times 1$ blocks on the diagonal of $T$, and the nonreal eigenvalues can be found as the eigenvalues of the $2 \times 2$ blocks on the diagonal of $T$.

The diagonal blocks of $T$ are called $T_i$. An important observation now is that each $T_i$ is weakly positive definite. This is trivially true if $T_i$ is a positive real number, and shown by Lemma 3.1 if $T_i$ has truly complex eigenvalues.

We now construct

$$B = QDQ^T,$$

[2]Standard Schur decomposition has an upper (not block) triangular matrix $T$, but is in general complex.

where $D$ is block diagonal, and where the diagonal blocks are denoted $D_i$. Then each $D_i$ should have the same dimension as $T_i$. We get that

$$BA = QDTQ^T.$$

In the following we show that the $D_i$s can be constructed in such a way that both $D$ and $DT$ are positive definite, and by Remark 2.3 this is equivalent to both $B$ and $BA$ being positive definite. Note that $D$ is positive definite if and only if all the $D_i$s are.

$T$ may be written

$$T = \begin{pmatrix} T_1 & \multicolumn{3}{c}{\boxed{\quad t_1 \quad}} \\ 0 & T_2 & \multicolumn{2}{c}{\boxed{\quad t_2 \quad}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T_k \end{pmatrix}$$

where $t_i$ denote the nonzero offdiagonal (usually nonsquare) blocks, i.e. $t_i$ is the matrix consisting of all the entries to the right of $T_i$. We see that

$$DT = \begin{pmatrix} D_1 & 0 & \dots & 0 \\ 0 & D_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D_k \end{pmatrix} \begin{pmatrix} T_1 & \multicolumn{3}{c}{\boxed{\quad t_1 \quad}} \\ 0 & T_2 & \multicolumn{2}{c}{\boxed{\quad t_2 \quad}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T_k \end{pmatrix}$$

$$= \begin{pmatrix} D_1T_1 & \multicolumn{3}{c}{\boxed{\quad D_1t_1 \quad}} \\ 0 & D_2T_2 & \multicolumn{2}{c}{\boxed{\quad D_2t_2 \quad}} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & D_kT_k \end{pmatrix}. \tag{6}$$

To finish this proof, we start in the lower right corner of $DT$ and show by induction that positive definite $D_i$ blocks can be constructed to make the block matrix (6) positive definite.

Since $T_k$ is weakly positive definite, there exists a positive definite $D_k$ such that the lower right block $D_kT_k$ is positive definite.

Similarly, there exists a positve definite $\tilde{D}_{k-1}$ such that $\tilde{D}_{k-1}T_{k-1}$ is positive definite. Choosing $D_{k-1} = \epsilon\tilde{D}_{k-1}$, Lemma 3.2 gives us that the extended lower right block

$$\begin{pmatrix} \epsilon\tilde{D}_{k-1}T_{k-1} & \epsilon\tilde{D}_{k-1}t_{k-1} \\ 0 & D_kT_k \end{pmatrix} \tag{7}$$

is positive definite (for any $t_{k-1}$) if $\epsilon$ is small enough.

By Lemma 3.2 and induction we see that the system (7) can be extended to the full block system (6), and the proof is complete. ☐

**Acknowledgements.** I would like to thank Sverre Lunøe–Nielsen, for interesting discussions – especially about $2 \times 2$ matrices. I am also thankful to Gilbert Strang who suggested Lemma 2.1. I will further thank Kent–Andre Mardal, Beresford Parlett, Gunnar Andreas Staff and Sigmund Vik for interesting discussions and comments. The interest for weakly positive definite matrices arose during discussions with Mardal.

## References

[1] Lloyd N. Trefethen and David Bau III, Numerical Linear Algebra, Siam, 1997.
[2] Gilbert Strang, Linear Algebra and Its Applications, 3rd ed., Harcourt, Brace, and Jovanovich, San Diego, 1988.
[3] K.-A. Mardal, T. K. Nilssen and G. A. Staff, Order optimal preconditioners for implicit Runge-Kutta schemes applied to parabolic PDEs, Research Report 08–2005, Simula Research Laboratory.
[4] T. K. Nilssen, K.-A. Mardal and G. T. Lines, Order optimal preconditioners for implicit Runge-Kutta schemes applied to the Bidomain equation, Research Report 09–2005, Simula Research Laboratory.

Simula Research Laboratory, Martin Linges vei 17, 1325 Lysaker, Norway.
*E-mail address*: trygvekn@simula.no

Now employed at Scandpower Petroleum Technology AS, P.O.Box 113, Gåsevikvn. 2-4, N-2027 Kjeller, Norway.
*E-mail address*: tkn@scandpowerpt.no