# Complex Cholesky-Jacobi Algorithm for PGEP

Vjeran Hari[1,a)]

[1]*Department of Mathematics, Faculty of Science, University of Zagreb, Bijenička cesta 30, 10000 Zagreb, Croatia.*

[a)]hari@math.hr
URL: http://web.math.hr/˜ hari

**Abstract.** In this report, we derive a new diagonalization method for the generalized eigenvalue problem $Ax = \lambda Bx$ with complex Hermitian $A$ and positive definite $B$. The method is a proper generalization to complex matrices of the real Cholesky-Jacobi method that was introduced in [1]. The method seems to be high relative accurate on pairs of well-behaved positive definite matrices.

## INTRODUCTION

The solutions of the generalized eigenvalue problem (GEP) $Ax = \lambda Bx$ with Hermitian matrix $A$ and Hermitian positive definite $B$ often appear in the course of solving practical problems of applied mathematics. Such problem is called positive definite GEP or shorter PGEP.

If the matrices $A$ and $B$ have large dimension, the best methods for solving the problem on contemporary CPU and GPU parallel computing machines are block Jacobi methods [3]. The block methods are designed to efficiently use the cache memory hierarchy and other features of modern computers. Typically, if the dimension of $A$ and $B$ is larger than 5000, then one block-column will include 32 or 64 adjacent columns. The block Jacobi methods should utilize accurate and fast kernel algorithms for solving the GEP with pivots submatrices, which are in our example of order 64 or 128. For that purpose, it is best to employ some element-wise Jacobi method because those methods are very accurate. On nearly diagonal matrices they are extremely fast and accurate.

For the latest problem, PGEP with complex matrices of small or moderate dimension, we propose a new method that we call Cholesky-Jacobi or shorter $CJ$ method. It employs the Cholesky factorization of the pivot submatrix of $B$ followed by the Jacobi step which diagonalizes the updated pivot submatrix of $A$. In [1] that method was derived for real symmetric matrices and here we derive it for complex Hermitian matrices. The numerical tests from [1] revealed excellent numerical properties of the real $CJ$ method which included the global convergence and high relative accuracy on pairs of well-behaved positive definite matrices. A positive definite matrix $H$ is well-behaved if there is a real diagonal matrix $D$ such that $DHD$ has the small spectral condition number. The method presupposes that the matrix $B$ has unit diagonal. This can be accomplished by the initial congruence transformation $DAD$, $DBD$, with the real diagonal matrix $D = [\mathrm{diag}(B)]^{-1/2}$. The $CJ$ method maintains the unit diagonal of $B$.

In this short paper, we are limited to only the derivation of one step of the element-wise $CJ$ method. The $2 \times 2$ pivot submatrices of $A$ and $B$ are denoted by $\hat{A}$ and $\hat{B}$ and their elements are subscripted by $i$ and $j$. Here, $(i, j)$, $i < j$ stands for the pivot pair which determines which rows and columns of the current iteration matrices will be transformed.

## THE DERIVATION OF THE COMPLEX $CJ$ ALGORITHM

The $CJ$ algorithm is comprised of two algorithms, the $LL^*J$ and $RR^*J$ algorithm. In each step the $CJ$ method employs the algorithm which is for the given input data more accurate.

## The $LL^*J$ Algorithm

Let us write the Cholesky factorization of $\hat{B}$ by elements,

$$\begin{bmatrix} 1 & b_{ij} \\ \bar{b}_{ij} & 1 \end{bmatrix} = \hat{B} = \hat{L}\hat{L}^* = \begin{bmatrix} 1 & 0 \\ \bar{a} & \bar{c} \end{bmatrix}\begin{bmatrix} 1 & a \\ 0 & c \end{bmatrix} = \begin{bmatrix} 1 & \bar{a} \\ a & |a|^2 + |c|^2 \end{bmatrix},$$

where $\bar{a}$ and $\hat{L}^*$ denote the complex conjugate $a$ and Hermitian transpose of $\hat{L}$. Assuming the real positive $c$, one immediately obtains $a = b_{ij}$, $c = \sqrt{1 - |b_{ij}|^2}$. Hence, using the notation $\tau = \sqrt{1 - |b_{ij}|^2}$, we obtain

$$\hat{L} = \begin{bmatrix} 1 & 0 \\ \bar{b}_{ij} & \tau \end{bmatrix}, \qquad \hat{L}^{-1} = \frac{1}{\tau}\begin{bmatrix} \tau & 0 \\ -\bar{b}_{ij} & 1 \end{bmatrix}, \qquad \hat{L}^{-*} = \frac{1}{\tau}\begin{bmatrix} \tau & -b_{ij} \\ 0 & 1 \end{bmatrix}.$$

If we write $\hat{F}_1 = \hat{L}^{-*}$, we obtain $\hat{F}_1^*\hat{B}\hat{F}_1 = I_2$. We also have

$$\begin{aligned} \hat{F}_1^*\hat{A}\hat{F}_1 &= \frac{1}{\tau^2}\begin{bmatrix} \tau & 0 \\ -\bar{b}_{ij} & 1 \end{bmatrix}\begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix}\begin{bmatrix} \tau & -b_{ij} \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} a_{ii} & (a_{ij} - b_{ij}a_{ii})/\sqrt{1 - |b_{ij}|^2} \\ (\bar{a}_{ij} - \bar{b}_{ij}a_{ii})/\sqrt{1 - |b_{ij}|^2} & a_{jj} - \frac{a_{ij}\bar{b}_{ij} + \bar{a}_{ij}b_{ij} - (a_{ii} + a_{jj})|b_{ij}|^2}{1 - |b_{ij}|^2} \end{bmatrix}. \end{aligned}$$

The final transformation $\hat{F}$ has the form $\hat{F} = \hat{F}_1\hat{R}_1$, where $\hat{R}_1$ is the complex Jacobi rotation that annihilates the off-diagonal element of $\hat{F}_1^*\hat{A}\hat{F}_1$. Let us assume that the $(1,2)$-element of $\hat{R}_1$ is $-e^{\imath\epsilon_1}\sin\vartheta_1$. Then the angles $\vartheta_1$ and $\epsilon_1$ are determined by the formulas

$$\begin{aligned} \epsilon_1 &= \arg(a_{ij} - b_{ij}a_{ii}), \\ \tan(2\vartheta_1) &= \frac{2|a_{ij} - a_{ii}b_{ij}|\sqrt{1 - |b_{ij}|^2}}{a_{ii} - a_{jj} + a_{ij}\bar{b}_{ij} + \bar{a}_{ij}b_{ij} - 2a_{ii}|b_{ij}|^2}, \quad -\frac{\pi}{4} \le \vartheta_1 \le \frac{\pi}{4}. \end{aligned}$$

The transformation formulas for the diagonal elements of $A$ read

$$\begin{aligned} a'_{ii} &= a_{ii} + \tan\vartheta_1 \cdot \frac{|a_{ij} - a_{ii}b_{ij}|}{\sqrt{1 - |b_{ij}|^2}}, \\ a'_{jj} &= a_{jj} - \frac{a_{ij}\bar{b}_{ij} + \bar{a}_{ij}b_{ij} - (a_{ii} + a_{jj})|b_{ij}|^2}{1 - |b_{ij}|^2} - \tan\vartheta_1 \cdot \frac{|a_{ij} - a_{ii}b_{ij}|}{\sqrt{1 - |b_{ij}|^2}}. \end{aligned}$$

In the case $a_{ii} = a_{jj}$, $a_{ij} = a_{ii}b_{ij}$ the expression for $\tan(2\vartheta_1)$ has the form $0/0$, and then we choose $\vartheta_1 = 0$. In that case $\hat{F}_1^*\hat{A}\hat{F}_1 = a_{ii}I_2$, hence $a'_{ii}$ and $a'_{jj}$ are reduced to $a_{ii}$ and $a_{jj}$, respectively.

Let $c_{\vartheta_1} = \cos\vartheta_1$, $s^{\pm}_{\vartheta_1} = e^{\pm\imath\epsilon_1}\sin\vartheta_1$. The transformation matrix $\hat{F}$ is obtained as follows.

$$\hat{F} = \frac{1}{\sqrt{1 - |b_{ij}|^2}}\begin{bmatrix} \sqrt{1 - |b_{ij}|^2} & -b_{ij} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} c_{\vartheta_1} & -s^{+}_{\vartheta_1} \\ s^{-}_{\vartheta_1} & c_{\vartheta_1} \end{bmatrix} = \frac{1}{\sqrt{1 - |b_{ij}|^2}}\begin{bmatrix} c_{\tilde{\vartheta}_1} & -s_{\tilde{\vartheta}_1} \\ s^{-}_{\vartheta_1} & c_{\vartheta_1} \end{bmatrix} = \begin{bmatrix} c1 & -s1 \\ s2 & c2 \end{bmatrix},$$

where

$$\begin{aligned} c_{\tilde{\vartheta}_1} &= c_{\vartheta_1}\sqrt{1 - |b_{ij}|^2} - s^{-}_{\vartheta_1}b_{ij}, \qquad & c1 &= c_{\vartheta_1} - s^{-}_{\vartheta_1}b_{ij}/\sqrt{1 - |b_{ij}|^2}, \qquad & c2 &= c_{\vartheta_1}/\sqrt{1 - |b_{ij}|^2}, \\ s_{\tilde{\vartheta}_1} &= c_{\vartheta_1}b_{ij} + s^{+}_{\vartheta_1}\sqrt{1 - |b_{ij}|^2}, & s1 &= c_{\vartheta_1}b_{ij}/\sqrt{1 - |b_{ij}|^2} + s^{+}_{\vartheta_1}, & s2 &= s^{-}_{\vartheta_1}/\sqrt{1 - |b_{ij}|^2}. \end{aligned}$$

It is easy to verify that $|c_{\tilde{\vartheta}_1}|^2 + |s_{\tilde{\vartheta}_1}|^2 = 1$. This algorithm works well, but we can still reduce the number of floating point operations per iteration step. This is acomplished by transforming the complex element $c1$ into $|c1|$. Formaly, we postmultiply $\hat{F}$ by the diagonal matrix $\mathrm{diag}(\bar{c}_{\tilde{\vartheta}_1}/|c_{\tilde{\vartheta}_1}|,\ 1)$, provided that $c_{\tilde{\vartheta}_1} \neq 0$. That transforms $s2$ into $s2 \cdot \bar{c}_{\tilde{\vartheta}_1}/|c_{\tilde{\vartheta}_1}|$.

## The $RR^*J$ Algorithm

Instead of $LL^*$, one can use $RR^*$ factorization of $\hat{B}$. Then we have

$$\begin{bmatrix} 1 & b_{ij} \\ \bar{b}_{ij} & 1 \end{bmatrix} = \hat{B} = \hat{R}\hat{R}^* = \begin{bmatrix} c & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{c} & 0 \\ \bar{a} & 1 \end{bmatrix} = \begin{bmatrix} |a|^2 + |c|^2 & a \\ \bar{a} & 1 \end{bmatrix}.$$

Assuming positive $c$, one obtains $a = b_{ij},\ c = \sqrt{1 - |b_{ij}|^2} = \tau$. Hence

$$\hat{R} = \begin{bmatrix} \tau & b_{ij} \\ 0 & 1 \end{bmatrix}, \qquad \hat{R}^{-1} = \frac{1}{\tau}\begin{bmatrix} 1 & -b_{ij} \\ 0 & \tau \end{bmatrix}, \qquad \hat{R}^{-*} = \frac{1}{\tau}\begin{bmatrix} 1 & 0 \\ -\bar{b}_{ij} & \tau \end{bmatrix}.$$

If we write $\hat{F}_2 = \hat{R}^{-*}$, then $\hat{F}_2^* \hat{B} \hat{F}_2 = \hat{R}^{-1}\hat{B}\hat{R}^{-*} = I_2$ and we have

$$\begin{aligned} \hat{F}_2^* \hat{A} \hat{F}_2 &= \frac{1}{\tau^2}\begin{bmatrix} 1 & -b_{ij} \\ 0 & \tau \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ \bar{a}_{ij} & a_{jj} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\bar{b}_{ij} & \tau \end{bmatrix} \\ &= \begin{bmatrix} a_{ii} - \frac{a_{ij}\bar{b}_{ij} + \bar{a}_{ij}b_{ij} - (a_{ii}+a_{jj})|b_{ij}|^2}{1-|b_{ij}|^2} & (a_{ij} - a_{jj}b_{ij})/\sqrt{1-|b_{ij}|^2} \\ (\bar{a}_{ij} - a_{jj}\bar{b}_{ij})/\sqrt{1-|b_{ij}|^2} & a_{jj} \end{bmatrix}. \end{aligned} \tag{1}$$

The final transformation $\hat{F}$ has the form $\hat{F} = \hat{F}_2\hat{R}_2$, where $\hat{R}_2$ is the Jacobi transformation which annihilates the off-diagonal element of $\hat{F}_2^*\hat{A}\hat{F}_2$. Let us assume that the $(1,2)$-element of $\hat{R}_2$ is $-e^{\imath\epsilon_2}\sin\vartheta_2$. Then the angles $\epsilon_2$ and $\vartheta_2$ are determined by the formulas

$$\begin{aligned} \epsilon_2 &= \arg(a_{ij} - b_{ij}a_{jj}), \\ \tan(2\vartheta_2) &= \frac{2|a_{ij} - a_{jj}b_{ij}|\sqrt{1-|b_{ij}|^2}}{a_{ii} - a_{jj} - (a_{ij}\bar{b}_{ij} + \bar{a}_{ij}b_{ij}) + 2a_{jj}|b_{ij}|^2}, \quad -\frac{\pi}{4} \le \vartheta_2 \le \frac{\pi}{4}. \end{aligned}$$

The transformation formulas for the diagonal elements of $A$ read

$$\begin{aligned} a'_{ii} &= a_{ii} - \frac{a_{ij}\bar{b}_{ij} + \bar{a}_{ij}b_{ij} - (a_{ii}+a_{jj})|b_{ij}|^2}{1-|b_{ij}|^2} + \tan\vartheta_2 \cdot \frac{|a_{ij} - a_{jj}b_{ij}|}{\sqrt{1-|b_{ij}|^2}}, \\ a'_{jj} &= a_{jj} - \tan\vartheta_2 \cdot \frac{|a_{ij} - a_{jj}b_{ij}|}{\sqrt{1-|b_{ij}|^2}}. \end{aligned}$$

In the case $a_{ii} = a_{jj}$, $a_{ij} = a_{jj}b_{ij}$ the angle $\vartheta_2$ is not well defined and then we choose $\vartheta_2 = 0$. In that case $a'_{ii}$ and $a'_{jj}$ are read from the relation (1) and they reduce to $a_{ii}$ and $a_{jj}$, respectively.

Let $c_{\vartheta_2} = \cos\vartheta_2,\ s^{\pm}_{\vartheta_2} = e^{\pm\imath\epsilon_2}\sin\vartheta_2$. Then the pivot submatrix $\hat{F}$ in the $RR^*J$ algorithm has the following form.

$$\hat{F} = \frac{1}{\sqrt{1-|b_{ij}|^2}}\begin{bmatrix} 1 & 0 \\ -\bar{b}_{ij} & \sqrt{1-|b_{ij}|^2} \end{bmatrix}\begin{bmatrix} c_{\vartheta_2} & -s^+_{\vartheta_2} \\ s^-_{\vartheta_2} & c_{\vartheta_2} \end{bmatrix} = \frac{1}{\sqrt{1-|b_{ij}|^2}}\begin{bmatrix} c_{\vartheta_2} & -s^+_{\vartheta_2} \\ s_{\tilde{\vartheta}_2} & c_{\tilde{\vartheta}_2} \end{bmatrix} = \begin{bmatrix} c1 & -s1 \\ s2 & c2 \end{bmatrix},$$

where

$$\begin{aligned} c_{\tilde{\vartheta}_2} &= c_{\vartheta_2}\sqrt{1-|b_{ij}|^2} + s^+_{\vartheta_2}\bar{b}_{ij}, &\quad c1 &= c_{\vartheta_2}/\sqrt{1-|b_{ij}|^2}, &\quad c2 &= c_{\vartheta_2} + s^+_{\vartheta_2}\bar{b}_{ij}/\sqrt{1-|b_{ij}|^2}, \\ s_{\tilde{\vartheta}_2} &= s^-_{\vartheta_2}\sqrt{1-b_{ij}^2} - c_{\vartheta_2}\bar{b}_{ij}, &\quad s1 &= s^+_{\vartheta_2}/\sqrt{1-|b_{ij}|^{2+}}, &\quad s2 &= s^-_{\vartheta_2} - c_{\vartheta_2}\bar{b}_{ij}/\sqrt{1-|b_{ij}|^2}. \end{aligned}$$

It is easy to show that $|c_{\tilde{\vartheta}_2}|^2 + |s_{\tilde{\vartheta}_2}|^2 = 1$. Again, we can postmultiply $\hat{F}$ by the diagonal matrix $\mathrm{diag}(1\,,\,\bar{c}_{\tilde{\vartheta}_2}/|c_{\tilde{\vartheta}_2}|)$ provided that $c_{\tilde{\vartheta}_2} \neq 0$. This ensures that (the updated) $\hat{F}$ has nonnegative diagonal elements.

## The $CJ$ Algorithm

The Cholesky-Jacobi is a hybrid algorithm which can be briefly defined as follows (cf. [1]): **select the pivot pair** $(i, j)$ **and if** $a_{ii} \le a_{jj}$ **then employ the** $LL^*J$ **algorithm, otherwise employ the** $RR^*J$ **algorithm.**

Our numerical tests show that neither $LL^*J$ nor $RR^*J$ is indicated as a high relative accurate algorithm on pairs of well-behaved positive definite matrices. The same can be said for the hybrid algorithm that selects the $LL^*J$ and $RR^*J$ algorithms in the opposite way, i.e. selects the $RR^*J$ ($LL^*J$) algorithm when $a_{ii} \leq a_{jj}$ ($a_{ii} > a_{jj}$). By the numerical tests, only the above definition seems to warrant the high relative accuracy of the algorithm. It is in complete agreement with the behavior of the real $LL^TJ$, $RR^TJ$ and $CJ$ methods from [1]. To solve the global convergence problem of the method, one can use the approach from [1] which uses the result from [2].

The first draft of the detailed pseudo code of the $CJ$ algorithm is given below. The logical variable *eivec* determines whether the matrix of eigenvectors $F = (f_{rt})$ will be computed. More implementation details, especially those addressing the stopping criterion of the method, will be published elsewhere.

**%%% Complex $CJ$ Algorithm**

```
select the pivot pair (i, j)
if a_ij ≠ 0 or b_ij ≠ 0 then
   β = |b_ij|;  ccb = conj(b_ij),  τ = sqrt((1 + β) * (1 − β));
   if  a_ii ≤ a_jj  then  σ = 1;  α1 = a_ii;  α2 = a_jj;
                  else  σ = −1;  α1 = a_jj;  α2 = a_ii;
   end;   e = a_ij − α1 * b_ij;
   if  e = 0  then  ee = 0;  ea = 1;  cs = 1,  sn = 0;  t = 0;
              else  ee = abs(e);  ea = e/ee;  ct2 = (0.5 * (α1 − α2) + Re(a_ij * ccb − α1 * β^2)/(σ * ee * τ);
                    t = sign(ct2)/(abs(ct2) + sqrt(1 + ct2^2));  cs = 1/sqrt(1 + t^2);  sn = t/sqrt(1 + t^2);
   end
   δ1 = σ * t * ee/τ;  δ2 = δ1 + (2 * Re(a_ij * ccb) − (α1 + α2) * β^2)/((1 − β) * (1 + β));  α1 = α1 + δ1;  α2 = α2 − δ2;
   if  σ > 0  then  c2 = cs/τ;  s2 = (sn/τ) * conj(ea);  c1 = cs − s2 * b_ij;  s1 = sn * ea + c2 * b_ij;
                      x = abs(c1);  s2 = s2 * conj(c1)/x;  c1 = x;  a'_ii = α1;  a'_jj = α2;
              else  c1 = cs/τ;  s1 = (sn/τ) * ea;  c2 = cs + s1 * ccb;  s2 = sn * conj(ea) − c1 * cbb;
                      x = abs(c2);  s1 = s1 * (cs + conj(s1) * b_ij)/x;  c2 = x;  a'_ii = α2;  a'_jj = α1;
   end;   s2c = conj(s2);
   a'_ij = (c1 * c2 * a_ij − s1 * s2c * conj(a_ij)) + (c2 * s2c * a_jj − c1 * s1 * a_ii);   a'_ji = conj(a'_ij);
   b'_ij = (c1 * c2 * b_ij − s1 * s2c * ccb + (c2 * s2c − c1 * s1);   b'_ji = conj(b'_ij);
   for  k = 1, ..., n,  k ≠ i, j
      a'_ki = c1 * a_ki + s2 * a_kj;   b'_ki = c1 * b_ki + s2 * b_kj;   a'_ik = conj(a'_ki);   b'_ik = conj(b'_ki);
      a'_kj = c2 * a_kj − s1 * a_ki;   b'_kj = c2 * b_kj − s1 * b_ki;   a'_jk = conj(a'_kj);   b'_jk = conj(b'_kj);
   end
   if eivec then,  for k = 1, ..., n,  f'_ki = c1 * f_ki + s2 * f_kj;  f'_kj = c2 * f_kj − s1 * f_ki;  end,  end
end
```


## ACKNOWLEDGMENTS

This work has been fully supported by Croatian Science Foundation under the project: IP_09_2014_3670. The author is thankful to the anonymous referees for the valuable comments.


## REFERENCES


[1] V. Hari, *Globally convergent Jacobi methods for positive definite matrix pairs*, Numer Algor (2017). https://doi.org/10.1007/s11075-017-0435-5.

[2] V. Hari, E. Begović Kovač, *Convergence of the Cyclic and Quasi-cyclic Block Jacobi Methods*. Electron. T. Numer. Ana. (ETNA), 46 (2017) 107-147.

[3] V. Novaković, S. Singer, S. Singer, *Blocking and Parallelization of the Hari-Zimmermann Variant of the Falk-Langemeyer Algorithm for the Generalized SVD*, Parallel Comput., 49 (2015) 136-152.