# Computing the 2-by-1 CS decomposition

Brian Sutton Randolph-Macon College

13 Aug 2012

This material is based upon work supported by the National Science Foundation under Grant No. DMS-0914559.

### Three decompositions:

2-by-2 CS decomposition (CSD)

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C - S \\ S C \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T$$
orthogonal
orthogonal
hlocks

2-by-1 CSD

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} & = & \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} & \begin{bmatrix} C \\ S \end{bmatrix} & V_1^T$$
 orthogonal diagonal orthogonal blocks

Generalized singular value decomposition (GSVD)

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} W^T$$
orthogonal diagonal blocks

Earlier work [Sutton 2009, 2012] emphasized the 2-by-2 CSD. This talk moves to the 2-by-1 CSD.

A CSD code must make *consistent* choices even when those choices are *arbitrary*.

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T$$

- The columns of  $U_1$  are the left singular vectors of  $X_{11}$  and  $X_{12}$ .
- The columns of  $U_2$  are the left singular vectors of  $X_{21}$  and  $X_{22}$ .
- The columns of  $V_1$  are the right singular vectors of  $X_{11}$  and  $X_{21}$ .
- The columns of  $V_2$  are the right singular vectors of  $X_{12}$  and  $X_{22}$ .
- Linear Algebra 101:  $\lambda_i = \lambda_{i+1} \implies$  two-dimensional eigen*space*.
- Numerical LA:  $\lambda_i \approx \lambda_{i+1} \implies ill$ -conditioned eigenvectors.
- CSD requires four simultaneous SVD's.
- The computed singular vectors must be identical.

Which is easier to compute?

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T$$

or

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V_1^T$$

## 2-by-1 CSD?

- Earlier algorithms (Stewart '82, Van Loan '85, Paige '86, Bai-Demmel '93, Drmač '98) compute the 2-by-1 CSD.
- The 2-by-1 CSD involves less sharing of singular vectors.

## 2-by-2 CSD?

 In "Computing the Complete CS Decomposition" (2009), I argue that the 2-by-2 CSD is easier....

Contrasted with the 2-by-1 CSD

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V_1^T,$$

the input to the 2-by-2 CSD

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T$$

- provides more information— $X_{12}$ ,  $X_{22}$ ,
- is more constrained—columns and rows are orthonormal.

Computing the 2-by-2 CSD ...

[	0.58	0.15	0.19	0.15	0.57	-0.45	-0.21	0.14	
	0.21	0.66	-0.14	-0.10	-0.34	-0.18	-0.05	-0.58	
	-0.53	0.68	0.23	0.01	0.19	0.05	-0.08	0.40	
l	0.32	0.11	0.04	-0.82	0.16	0.34	0.23	0.15	
	-0.43	-0.13	-0.03	-0.26	0.44	-0.47	0.40	-0.40	
	-0.05	0.01	-0.74	-0.22	-0.14	-0.42	-0.14	0.43	
١	-0.15	-0.22	0.49	-0.43	-0.31	-0.34	-0.54	-0.01	
	0.16	-0.04	-0.33	-0.09	0.44	0.37	-0.64	-0.33	

= X

0.21	0.66	-0.14	-0.10	-0.34	-0.18	-0.05	-0.58
-0.53	0.68	0.23	0.01	0.19	0.05	-0.08	0.40
0.32	0.11	0.04	-0.82	0.16	0.34	0.23	0.15
-0.43	-0.13	-0.03	-0.26	0.44	-0.47	0.40	-0.40
-0.05	0.01	-0.74	-0.22	-0.14	-0.42	-0.14	0.43
-0.15	-0.22	0.49	-0.43	-0.31	-0.34	-0.54	-0.01
-0.16	-0.04	-0.33	-0.09	0.44	0.37	-0.64	-0.33

**0.58** 0.15 0.19 0.15 0.57 -0.45 -0.21 0.14

= X

$$=\begin{bmatrix} F \\ F \end{bmatrix} X$$

-0.31

-0.03

0.87

-0.11

0.01

0.03

0.29

0.24

-0.190.53

-0.25

-0.02

-0.71

-0.24

Γ	0.87	-0.11	-0.03	-0.23	0.24	-0.25	-0.02	-0.24
		-0.66	0.16	0.08	0.46	0.07	0.02	0.57
-		-0.70	-0.29	0.02	-0.48	0.20	0.16	-0.37
-		-0.10	-10 <sup>-3</sup>	0.80	0.01	-0.49	-0.28	-0.17
	0.49	0.19	0.06	0.41	-0.42	0.44	0.04	0.42
		0.03	-0.73	-0.18	-0.19	-0.37	-0.17	0.48
-		-0.16	0.50	-0.32	-0.45	-0.19	-0.60	0.12

$$=\begin{bmatrix} F \\ F \end{bmatrix} X$$

 $\begin{bmatrix} -0.16 & 0.50 & -0.32 & -0.45 & -0.19 & -0.60 & 0.12 \\ 0.01 & -0.31 & 0.03 & 0.29 & 0.53 & -0.71 & -0.19 \end{bmatrix}$ 

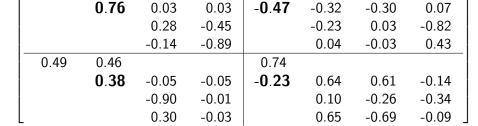
Γ	0.87	-0.26			-0.42			1
		-0.18	-0.24	-0.61	0.11	0.21	0.03	0.69
-		-0.31	0.19	-0.66	0.19	-0.06	0.14	-0.61
l		0.67	0.05	-0.44	-0.41	-0.33	-0.27	-0.02
	0.49	0.46			0.74			
		-0.24	0.71	0.04	0.15	-0.50	-0.18	0.35
		-0.29	-0.54	0.03	0.18	-0.43	-0.62	-0.11
L		-0.01	0.31	-0.03	0.01	0.64	-0.70	-0.08

$$=\begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix}$$

$$= \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix}$$

Γ	0.87	-0.26			-0.42			7
		-0.18	-0.24	-0.61	0.11	0.21	0.03	0.69
İ		-0.31	0.19	-0.66	0.19	-0.06	0.14	-0.61
		0.67	0.05	-0.44	-0.41	-0.33	-0.27	-0.02
	0.49	0.46			0.74			
		-0.24	0.71	0.04	0.15	-0.50	-0.18	0.35
		-0.29	-0.54	0.03	0.18	-0.43	-0.62	-0.11
L		-0.01	0.31	-0.03	0.01	0.64	-0.70	-0.08

$$= \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix}$$



$$= \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix}$$

0.87

0.03

0.03

-0.42

-0.47

-0.30

0.07

-0.82

0.43

-0.14

-0.34

-0.09

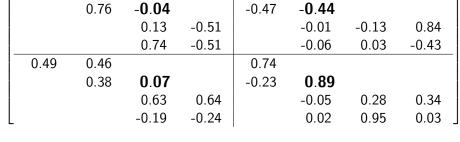
-0.32

0.87

-0.26

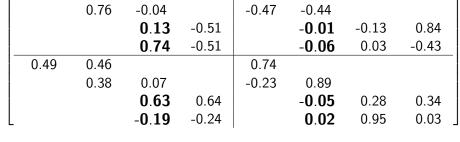
0.76

$$= \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix}$$



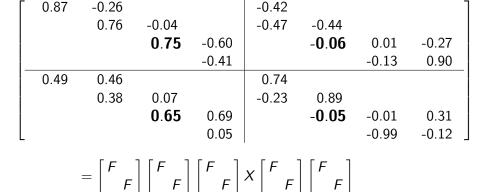
$$= \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix}$$

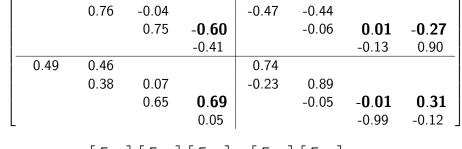
0.87



$$= \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix}$$

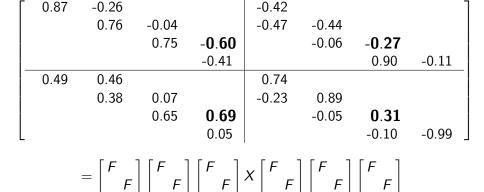
0.87

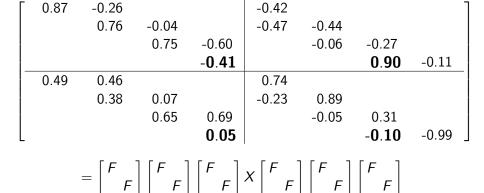


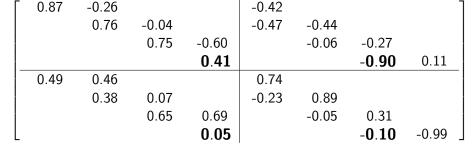


$$= \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix}$$

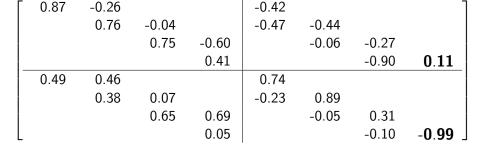
0.87







$$= \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix}$$



$$= \begin{bmatrix} F \\ F \end{bmatrix} X \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} F \\ F \end{bmatrix}$$

0.87

$$= \begin{bmatrix} F \\ F \end{bmatrix} \begin{bmatrix} I \\ F \end{bmatrix}$$

 $= \left[ \begin{array}{c|c} F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | & F & | &$ 

[Sutton (under review)].

In simultaneous bidiagonalization for the 2-by-2 CSD, a Householder reflector is

- constructed from two collinear columns or rows,
- applied to two collinear columns or rows.

The extra information in the 2-by-2 CSD is helpful—one of the two vectors is guaranteed to have norm  $\geq 1/\sqrt{2}$ , making the Householder reflector well determined.

For the 2-by-1 CSD 
$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V_1^T$$
,

- half of the input is, in a sense, missing,
- rows are not orthonormal.

Can we simultaneously bidiagonalize  $X_{11}$  and  $X_{21}$  stably?

For the 2-by-1 CSD 
$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V_1^T$$
,

- half of the input is, in a sense, missing,
- rows are not orthonormal.

Can we simultaneously bidiagonalize  $X_{11}$  and  $X_{21}$  stably?

### Yes!

• Full QR decomposition.

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

2-by-2 CSD.

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C & -S \\ S & C \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}^T$$

Discard the right half.

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V_1^T$$

Can we simultaneously bidiagonalize  $X_{11}$  and  $X_{21}$  stably and efficiently?

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} V_1^T$$
 orthogonal bidiagonal orthogonal blocks

This is our problem!

The key trick in simultaneous bidiagonalization is to construct a Householder reflector from two collinear columns or rows.

With the 2-by-1 CSD, we never have the second column.

Can it simply be ignored?

The key trick in simultaneous bidiagonalization is to construct a Householder reflector from two collinear columns or rows.

With the 2-by-1 CSD, we never have the second column.

Can it simply be ignored?

No!

$$\begin{bmatrix} 0.58 & 0.13 & -0.16 & 0.78 \\ -0.05 & -0.76 & 0.17 & 0.28 \\ -0.12 & -0.18 & 0.70 & 0.17 \\ -0.05 & -0.55 & -0.41 & -0.02 \\ \hline 0.68 & -0.24 & 0.10 & -0.44 \\ -0.22 & -0.01 & -0.20 & 0.11 \\ 0.36 & -0.02 & 0.12 & -0.25 \\ 0.02 & 0.16 & 0.47 & 0.11 \\ \hline \end{bmatrix}$$

 $\left\|I - X^T X\right\|_2 \approx 2.2 \times 10^{-15}$ 

「 0.58	0.13	-0.16	0.78
-0.05	-0.76	0.17	0.28
-0.12	-0.18	0.70	0.17
-0.05	-0.55	-0.41	-0.02
0.68	-0.24	0.10	-0.44
-0.22	-0.01	-0.20	0.11
0.36	-0.02	0.12	-0.25
0.02	0.16	0.47	0.11

0.59	0.27	-0.28	0.71
$10^{-18}$	0.74	-0.15	-0.34
$10^{-17}$	0.14	-0.66	-0.31
$10^{-18}$	0.53	0.43	-0.04
0.81	-0.20	0.20	-0.52
$10^{-17}$	0.08	0.16	0.03
$-10^{-17}$	-0.09	-0.05	0.02
<b>10</b> <sup>-18</sup>	-0.16	-0.46	-0.12

Γ 0.59	0.27	-0.28	0.71
10 <sup>-18</sup>	0.74	-0.15	-0.34
10 <sup>-17</sup>	0.14	-0.66	-0.31
10 <sup>-18</sup>	0.53	0.43	-0.04
0.81	-0.20	0.20	-0.52
10 <sup>-17</sup>	0.08	0.16	0.03
-10 <sup>-17</sup>	-0.09	-0.05	0.02
-10 <sup>-18</sup>	-0.16	-0.46	-0.12

Γ	0.59	-0.81	$10^{-16}$	- <b>10</b> <sup>-16</sup> ]
	$10^{-18}$	$-10^{-11}$	0.04	-0.83
	$10^{-17}$	$10^{-11}$	-0.62	-0.40
	$10^{-18}$	$10^{-11}$	0.56	-0.38
	0.81	0.59	-10 <sup>-17</sup>	0
	$10^{-17}$	-10 <sup>-11</sup>	0.18	-0.02
	-10 <sup>-17</sup>	$10^{-11}$	-0.07	0.08
L	-10 <sup>-18</sup>	-10 <sup>-11</sup>	-0.51	-0.01

0.59	-0.81	$10^{-16}$	-10 <sup>-16</sup>
$10^{-18}$	-10 <sup>-11</sup>	0.04	-0.83
$10^{-17}$	$10^{-11}$	-0.62	-0.40
$10^{-18}$	$10^{-11}$	0.56	-0.38
0.81	0.59	-10 <sup>-17</sup>	
$10^{-17}$	$-10^{-11}$	0.18	-0.02
$-10^{-17}$ $-10^{-18}$	$10^{-11}$	-0.07	0.08
-10 <sup>-18</sup>	- <b>10</b> <sup>-11</sup>	-0.51	-0.01

0.59	-0.81	$10^{-16}$	-10 <sup>-16</sup>
$10^{-18}$	$10^{-11}$	0.03	-0.13
$10^{-17}$	$10^{-27}$	-0.62	-0.67
$10^{-18}$	$10^{-27}$	0.57	-0.73
0.81	0.59	-10 <sup>-17</sup>	
-10 <sup>-16</sup>	$10^{-11}$	-0.02	0.08
$10^{-17}$	$10^{-26}$	0.07	0.01
-10 <sup>-17</sup>	-10 <sup>-27</sup>	-0.54	$10^{-3}$

0.59	-0.81	$10^{-16}$	-10 <sup>-16</sup>	
10 <sup>-18</sup>	10-11	0.03	-0.13	
$10^{-17}$	$10^{-27}$	-0.62	-0.67	
10 <sup>-18</sup>	$10^{-27}$	0.57	-0.73	
0.81	0.59	$-10^{-17}$		
-10 <sup>-16</sup>	10-11	-0.02	0.08	
10 <sup>-17</sup>	$10^{-26}$	0.07	0.01	
-10 <sup>-17</sup>	-10 <sup>-27</sup>	-0.54	$10^{-3}$	

10 <sup>-18</sup> 10 <sup>-17</sup>	-0.81 10 <sup>-11</sup> 10 <sup>-27</sup>	-10 <sup>-16</sup> - <b>0</b> . <b>13</b> -0.50	10 <sup>-17</sup> - <b>10</b> <sup>-6</sup> -0.76
10 <sup>-18</sup>	$10^{-27}$	-0.84	0.38
0.81	0.59	$10^{-18}$	$-10^{-17}$
-10 <sup>-16</sup>	$10^{-11}$	0.08	$-10^{-6}$
10 <sup>-17</sup>	$10^{-26}$	-0.01	0.07
L -10 <sup>-17</sup>	-10 <sup>-27</sup>	0.13	-0.52

0.59	-0.81	-10 <sup>-16</sup>	$10^{-17}$
$10^{-18}$	$10^{-11}$	-0.13	-10 <sup>-6</sup>
$10^{-17}$	$10^{-27}$	-0.50	-0.76
$10^{-18}$	$10^{-27}$	-0.84	0.38
0.81	0.59	10 <sup>-18</sup>	-10 <sup>-17</sup>
-10 <sup>-16</sup>	$10^{-11}$	0.08	-10 <sup>-6</sup>
$10^{-17}$	$10^{-26}$	-0.01	0.07
-10 <sup>-17</sup>	$-10^{-27}$	0.13	-0.52

$0.59 \\ 10^{-18} \\ -10^{-17} \\ -10^{-18} \\ \hline 0.81 \\ -10^{-16} \\ -10^{-17}$
---

Γ	0.59	-0.81	-10 <sup>-16</sup>	10 <sup>-17</sup>
	$10^{-18}$	10 <sup>-11</sup>	-0.13	-10 <sup>-6</sup>
	-10 <sup>-17</sup>	-10 <sup>-27</sup>	0.98	0.07
	-10 <sup>-18</sup>	-10 <sup>-27</sup>	$10^{-16}$	0.85
	0.81	0.59	$10^{-18}$	-10 <sup>-17</sup>
	-10 <sup>-16</sup>	$10^{-11}$	0.08	-10 <sup>-6</sup>
	$-10^{-17}$	-10 <sup>-27</sup>	0.13	-0.52
L	$10^{-17}$	$10^{-26}$	$10^{-18}$	0.03

Γ 0.59	-0.81	-10 <sup>-16</sup>	-10 <sup>-17</sup>
10 <sup>-18</sup>	$10^{-11}$	-0.13	10-6
-10 <sup>-17</sup>	-10 <sup>-27</sup>	0.98	-0.07
10 <sup>-18</sup>	-10 <sup>-27</sup>	10 <sup>-16</sup>	-0.85
0.81	0.59	10 <sup>-18</sup>	10 <sup>-17</sup>
-10 <sup>-16</sup>	$10^{-11}$	0.08	10 <sup>-6</sup>
-10 <sup>-17</sup>		0.13	0.52
L 10 <sup>-17</sup>	$10^{-26}$	$10^{-18}$	-0.03

0.59	-0.81	-10 <sup>-16</sup>	-10 <sup>-17</sup>
10 <sup>-18</sup>	10-11	-0.13	10 <sup>-6</sup>
-10 <sup>-17</sup>	-10 <sup>-27</sup>	0.98	-0.07
10 <sup>-18</sup>	-10 <sup>-27</sup>	$10^{-16}$	-0.85
0.81	0.59	$10^{-18}$	10 <sup>-17</sup>
-10 <sup>-16</sup>	$10^{-11}$	0.08	10 <sup>-6</sup>
-10 <sup>-17</sup>	-10 <sup>-27</sup>	0.13	0.52
L 10 <sup>-17</sup>	$10^{-26}$	$10^{-18}$	-0.03 📗

0.59	-0.81	-10 <sup>-16</sup>	-10 <sup>-17</sup>
10 <sup>-18</sup>	$10^{-11}$	-0.13	10 <sup>-6</sup>
-10 <sup>-17</sup>	-10 <sup>-27</sup>	0.98	-0.07
10 <sup>-18</sup>	10 <sup>-27</sup>	-10 <sup>-16</sup>	0.85
0.81	0.59	$10^{-18}$	10 <sup>-17</sup>
-10 <sup>-16</sup>	$10^{-11}$	0.08	10 <sup>-6</sup>
-10 <sup>-17</sup>	-10 <sup>-27</sup>	0.13	0.52
L -10 <sup>-17</sup>	-10 <sup>-26</sup>	-10 <sup>-18</sup>	0.03

0.59	-0.81		
	$10^{-11}$	-0.13	$10^{-6}$
		0.98	-0.07
			0.85
0.81	0.59		
	$10^{-11}$	0.08	$10^{-6}$
		0.13	0.52
-			0.03

Γ	0.59	-0.81		-
		$10^{-11}$	-0.13	$10^{-6}$
İ			0.98	-0.07
				0.85
[	0.81	0.59		
		$10^{-11}$	0.08	$10^{-6}$
			0.13	0.52
L				0.03

Some entries are not numerically zero as they should be.

This does not happen for the 2-by-2 case. The proof relies on the orthogonality of the rows in addition to the columns.

Can we simultaneously bidiagonalize  $X_{11}$  and  $X_{21}$  stably and efficiently?

Yes, but we need information from  $X_{12}$  and  $X_{22}$ .

Instead of generating the right half of the matrix explicitly with a QR factorization—

$$\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix}$$

—generate it implicitly, one column at a time, without using any extra storage.

×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
X	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
L ×	×	×	×	?	?	?	?

Problem: Simultaneously bidiagonalize the blocks on the left.

Method: Pretend that the right half of the matrix exists but has not been observed. Complete it one column at a time.

×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
X	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?
×	×	×	×	?	?	?	?

Householder reflectors from left...

c <sub>1</sub>	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
s <sub>1</sub>	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	? ]

$$c_i = \cos \theta_i, \qquad s_i = \sin \theta_i$$

$c_1$	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
$s_1$	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	7

Householder reflector from right; complete rows 1 and 5...

[ c	1	$-s_1s_1'$			$-s_1c_1'$			
		×	×	×	?	?	?	?
		×	×	×	?	?	?	?
ļ 		×	×	×	?	?	?	?
s	1	$c_1s_1'$			$c_1c_1'$			
		×	×	×	?	?	?	?
		×	×	×	?	?	?	?
1			×	×	?	2	2	?

 $c_i = \cos \theta_i, \qquad s_i = \sin \theta_i, \qquad c'_i = \cos \phi_i, \qquad s'_i = \sin \phi_i$ 

Rows 1 and 5 are orthonormal.

$c_1$	$-s_1s_1'$			$-s_1c_1'$			
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?
$s_1$	$c_1s_1'$			$c_1c_1'$			
	×	×	×	?	?	?	?
	×	×	×	?	?	?	?

Now for the tricky part. We would like to construct Householder reflectors from columns 2 and 5, but we don't have column 5.

The second column must have a specific form by orthogonality....

$c_1$	$-s_1s_1'$			$-s_1c_1'$			
	:	×	×	?	?	?	?
	$c_1'x$	×	×	?	?	?	?
	:	×	×	?	?	?	?
$s_1$	$c_1s_1'$			$c_1c_1'$			
	• • •	×	×	?	?	?	?
	$c_1'y$	×	×	?	?	?	?

Orthogonality forces ||(x, y)|| = 1.

Notice that we cannot measure (x,y) reliably; we only have  $c_1'(x,y)$ . Orthogonalize against columns 3, 4 using Kahan's "twice is enough." (If the projection is zero, then choose another vector arbitrarily.)

Then, complete column 5....

$c_1$	$-s_1s_1'$			$-s_1c_1'$			
	:	×	×	:	?	?	?
	$c_1' \bar{x}$	×	×	$-s_1'\bar{x}$	?	?	?
	:	×	×		?	?	?
$s_1$	$c_1s_1'$			$c_1c_1'$			
$s_1$	$c_1s_1'$	×	×	$c_1c_1'$	?	?	?
<i>S</i> <sub>1</sub>	$c_1s_1'$ $\vdots$ $c_1'ar{y}$	×	×	$\begin{vmatrix} c_1c_1' \\ \vdots \\ -s_1'\bar{y} \end{vmatrix}$	?	?	?

Columns  $1, \ldots, 5$  are orthonormal.

$c_1$	$-s_1s_1'$			$-s_1c_1'$			_
	•	×	×	:	?	?	?
	$c_1'\bar{x}$	×	×	$-s_1'\bar{x}$	?	?	?
	•	×	×	:	?	?	?
$s_1$	$c_1s_1'$			$c_1c_1'$			
	•	×	×	•	?	?	?
	$c_1'ar{y}$	×	×	$-s_1'ar{y}$	?	?	?
	•	×	×	*	?	?	?

$c_1$	$-s_1s_1'$			$-s_1c_1'$			-
	$c_2c_1'$	×	×	$-c_2s_1'$	?	?	?
		×	×		?	?	?
		×	×		?	?	?
$s_1$	$c_1s_1'$			$c_1c_1'$			
	$s_2c_1'$	×	×	$-s_2s_1'$	?	?	?
		×	×		?	?	?
		×	×		?	?	?

Continue....

	$c_1$	$-s_1s_1'$			$-s_1c_1'$
		$c_2 c_1'$	$-s_2s_2'$		$-c_2s_1' - s_2c_2'$
			$c_3c_2'$	$-s_3s_3'$	$-c_3s_2' -s_3c_3'$
ļ 				$c_4c_3'$	$-c_4s_3'$ $-s_4$
	$s_1$	$c_1s_1'$			$c_1c_1'$
		$s_2c_1'$	$c_2s_2'$		$-s_2s_1'$ $c_2c_2'$
			$s_3c_2'$	$c_3s_3'$	$-s_3s_2'$ $c_3c_3'$
				$s_4c_3'$	$-s_4s_3'$ $c_4$

We've actually computed

$$\begin{bmatrix} U_1 & & \\ & U_2 \end{bmatrix}^T \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} V_1 & \\ & I \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

with  $X_{12} := U_1 B_{12}$  and  $X_{22} := U_2 B_{22}$ .

This has been done without storing  $X_{12}$  or  $X_{22}$  in computer memory.

To obtain the 2-by-1 CSD, discard the right half:

$$\begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix}^T \begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix} V_1 = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}.$$

## Summary:

- If  $\begin{bmatrix} X_{11} \\ X_{21} \end{bmatrix}$  has orthonormal columns, then its blocks can be simultaneously bidiagonalized.
- The reduction is numerically stable.
  - The naive approach does not work.
  - The solution uses an extra orthogonalization step.
- The reduction is efficient.
  - Typically, the orthogonalization step is dominated by a single matrix-vector multiply. (If the numerical projection is zero, then the search for a vector in the orthogonal complement requires one or more additional matrix-vector multiplies.)

## References:

[Sutton 2009] "Computing the complete CS decomposition." *Numer. Algorithms.* 50 (2009), no. 1, 33-65.