Downdating and Other Operations on a Truncated Complete Orthogonal Decomposition

Jesse L. Barlow
The Pennsylvania State University
University Park, PA USA
barlow@cse.psu.edu

URL: `www.cse.psu.edu/~barlow/downdate_slides.ps`

The Problem

$$X \in \mathsf{R}^{m\times n}, m \geq n$$

Want to write

$$\begin{aligned} X &= U_1 L V_1^T + E \\ U_1^T U_1 &= V_1^T V_1 = I \quad \text{left orthogonal} \\ \|L^{-1}\|^{-1} &\geq \epsilon \geq \|E\|, \quad \text{Ideally} \\ U_1 &\in \mathsf{R}^{m\times k}, \quad V_1 \in \mathsf{R}^{n\times k} \\ L &\in \mathsf{R}^{k\times k}, \quad k \leq n \end{aligned}$$

Want to keep only

$$U_1, \quad L, \quad V_1$$

and the ability to do

$$X\mathbf{v}, \quad X^T\mathbf{u}$$

Useful if X is sparse or structured.

Needed Operations

Formation

truncated SVD
truncated ULVD
truncated bidiagonal reduction plus partial SVD

Updating

Add a row to X

Downdating

Delete a row from X

Norm Estimation

Compute estimates for $\|L^{-1}\|$ or $\|E\|$

Refinement

Make $\|L^{-1}\|$ or $\|E\|$ smaller

Previous Versions

ULVD

Stewart (1993) – Updating
Fadeev et al. (1968) –Formation
Hanson and Lawson (1969,1974) – Formation
Fierro and Bunch (1995) – Formation and subspaces

$$\begin{aligned} X &= UCV^T \\ &= \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} L & 0 \\ F & G \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} \\ &= U_1 L V_1^T + U_2 \begin{pmatrix} F & G \end{pmatrix} V^T \end{aligned}$$

Downdating and Norm Estimation
B. ,Yoon, & Zha (1996)
Yoon & B. (1998)
Erbay, B., and Zhang (2003)

Stored U_2, F,G, V_2 explicitly!

Necessary Tools

$$X = U_1 L V_1^T + E$$

Find $\mathbf{u}_k$,$\mathbf{v}_k$ s.t.

$$L\mathbf{v}_k = \sigma_k \mathbf{u}_k, \quad \sigma_k \approx \|L^{-1}\|^{-1}$$

Inverse Lanczos or power iterations

Find $\mathbf{u}_0$,$\mathbf{v}_0$ s.t.

$$E\mathbf{v}_0 = \sigma_1 \mathbf{u}_0, \quad \sigma_1 \approx \|E\|$$

Lanczos or power iterations on

$$(I - U_1 U_1^T)X$$

Wise in this context to insure that

$$U_1^T \mathbf{u}_0 = 0$$

to the extent possible.

Refinement

Mathias and Stewart (1995) – $O(mn^2)$ routine

Impractical for this setting

Our Procedure– When a row may be added.

Based on that by B., Erbay, and Slapničar (2003)

Find $\mathbf{u}_0$, $\mathbf{v}_0$ s.t.

$$E\mathbf{v}_0 = \sigma_1 \mathbf{u}_0$$

Compute

$$\mathbf{z} = X^T \mathbf{u}_0 (= E^T \mathbf{u}_0)$$

Find $\mathbf{v}_{k+1}$ (using Gram-Schmidt) s.t.

$$\begin{aligned} \mathbf{z} &= V_1 \mathbf{g} + \alpha \mathbf{v}_{k+1} \\ &= \begin{pmatrix} V_1 & \mathbf{v}_{k+1} \end{pmatrix} \begin{pmatrix} \mathbf{g} \\ \alpha \end{pmatrix} \\ V_1^T \mathbf{v}_{k+1} &= 0 \end{aligned}$$

Then

$$X = \begin{pmatrix} U_1 & \mathbf{u}_0 \end{pmatrix} \begin{pmatrix} L & 0 \\ \mathbf{g}^T & \alpha \end{pmatrix} \begin{pmatrix} V_1^T \\ \mathbf{v}_{k+1}^T \end{pmatrix} + \widehat{E}$$

Refinement (2)

Find $\widehat{\sigma}_{k+1}$ s.t.

$$\begin{aligned}
\begin{pmatrix} L & 0 \\ \mathbf{g}^T & \alpha \end{pmatrix} \mathbf{z}_{k+1} &= \widehat{\sigma}_{k+1} \mathbf{y}_{k+1} \\
\widehat{\sigma}_{k+1} &\approx \| \begin{pmatrix} L & 0 \\ \mathbf{g}^T & \alpha \end{pmatrix}^{-1} \|^{-1} \approx \sigma_{k+1}(X)
\end{aligned}$$

If $\widehat{\sigma}_{k+1} > \epsilon$ set

$$\begin{aligned}
L &= \begin{pmatrix} L & 0 \\ \mathbf{g}^T & \alpha \end{pmatrix} \\
U_1 &= \begin{pmatrix} U_1 & \mathbf{u}_0 \end{pmatrix}, \quad V_1 = \begin{pmatrix} V_1 & \mathbf{v}_{k+1} \end{pmatrix}
\end{aligned}$$

If $\widehat{\sigma}_{k+1} > \epsilon$ set

$$\begin{aligned}
L &= \begin{pmatrix} L & 0 \\ \mathbf{g}^T & \alpha \end{pmatrix} \\
U_1 &= \begin{pmatrix} U_1 & \mathbf{u}_0 \end{pmatrix}, \quad V_1 = \begin{pmatrix} V_1 & \mathbf{v}_{k+1} \end{pmatrix}
\end{aligned}$$

Refinement (3) – Else Part

Else let

$$Q^T \begin{pmatrix} L & 0 \\ \mathbf{g}^T & \alpha \end{pmatrix} Z = \begin{pmatrix} \widehat{L} & 0 \\ 0 & \widehat{\sigma}_{k+1} \end{pmatrix}$$

Take

$$\begin{aligned} \begin{pmatrix} \widehat{U}_1 & \widehat{\mathbf{u}}_0 \end{pmatrix} &= \begin{pmatrix} U_1 & \mathbf{u}_0 \end{pmatrix} Q \\ \begin{pmatrix} \widehat{V}_1 & \widehat{\mathbf{v}}_{k+1} \end{pmatrix} &= \begin{pmatrix} V_1 & \mathbf{v}_{k+1} \end{pmatrix} Z \\ L &= \widehat{L}, \quad U_1 = \widehat{U}_1 \quad V_1 = \widehat{V}_1 \\ \widehat{\sigma}_{k+1} &\approx \sigma_{k+1}(X) \\ \sigma_{k+1}^2(X) - \sigma_2^2(E) &\leq \widehat{\sigma}_{k+1}^2 \leq \sigma_{k+1}^2(X) \end{aligned}$$

Downdating

Yoon and B. (1998)
Erbay, B.,and Zhang (2003)

We have

$$X = \begin{pmatrix} \mathbf{x}_0^T \\ \bar{X} \end{pmatrix} = U_1 L V_1^T + E$$

to get

$$\bar{X} = \bar{U}_1 L \bar{V}_1^T + \bar{E}$$

Use Gram–Schmidt procedure from
B.,Smoktunowicz, and Erbay (2003) so that

$$\begin{aligned} \begin{pmatrix} \mathbf{u}_0 & U_1 \end{pmatrix} \mathbf{d} &= \mathbf{e}_1 \\ U_1^T \mathbf{u}_0 &= 0 \end{aligned}$$

Both should be satisfied as completely as possible

Daniel et al. (1976), Yoo and Park (1996)

Downdating (2)

Use same procedure find $\mathbf{g}$,α s.t.

$$\begin{aligned} X^T\mathbf{u}_0 &= V_1\mathbf{g} + \alpha\mathbf{v}_{k+1} \\ V_1^T\mathbf{v}_{k+1} &= 0 \end{aligned}$$

Then

$$\begin{pmatrix} \mathbf{x}_0^T \\ \bar{X} \end{pmatrix} = \begin{pmatrix} \mathbf{u}_0 & U_1 \end{pmatrix} \begin{pmatrix} \mathbf{g}^T & \alpha \\ L & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ \mathbf{v}_{k+1}^T \end{pmatrix} + \tilde{E}$$

No components of $\tilde{E}$ in the first row.

Find Q, Z orthogonal s.t.

$$\begin{aligned} Q^T\mathbf{d} &= \|\mathbf{d}\|\mathbf{e}_1, \quad \|\mathbf{d}\| \approx 1 \\ Q^T \begin{pmatrix} \mathbf{g}^T & \alpha \\ L & 0 \end{pmatrix} Z &= \begin{pmatrix} \tilde{\mathbf{g}}^T & \tilde{\alpha} \\ \tilde{L} & 0 \end{pmatrix} \end{aligned}$$

Downdating (3)

$$\begin{pmatrix} \mathbf{u}_0 & U_1 \end{pmatrix} Q = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U}_1 \end{pmatrix}$$

"In theory"

"In practice" – see B., Smoktunowicz, and Erbay (2003)

$$\begin{pmatrix} V_1 & \mathbf{v}_{k+1} \end{pmatrix} Z = \begin{pmatrix} \tilde{V}_1 & \tilde{\mathbf{v}}_{k+1} \end{pmatrix}$$

Then

$$\begin{pmatrix} \mathbf{x}_0^T \\ \bar{X} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{U}_1 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{g}}^T & \tilde{\alpha} \\ \tilde{L} & 0 \end{pmatrix} \begin{pmatrix} \tilde{V}_1^T \\ \tilde{\mathbf{v}}_{k+1}^T \end{pmatrix} + \tilde{E}$$

or

$$\bar{X} = \tilde{U}_1 \tilde{L} \tilde{V}_1 + \tilde{E}(2{:}\, m, :)$$

Downdating (4)

We know

$$rank(L) - 1 \leq rank(\tilde{L}) \leq rank(L)$$

Find

$$\tilde{L}\mathbf{v}_k = \sigma_k \mathbf{u}_k$$

If $\sigma_k > \epsilon$

$$\bar{U}_1 = \tilde{U}_1, \quad \bar{L} = \tilde{L}, \quad \bar{V}_1 = \tilde{V}_1$$

else find Q_1,Z_1 s.t.

$$Q_1^T \tilde{L} Z_1 = \begin{pmatrix} \bar{L} & 0 \\ 0 & \sigma_k \end{pmatrix}$$

Norm Bounding

Good News $\|E\|_F$ and $\|L^{-1}\|_F$ can be updated exactly in $O(k)$ and $O(k^2)$ flops respectively.

Bad News These bounds are not sharp enough and these algorithms are not stable.

However,

$$\|E\|_F^2 = \|X\|_F^2 - \|L\|_F^2$$

requires $O(mn)$ flops or less. Still not sharp enough!

Two Norm Bound

Maintain

$$bound(E) \geq \|E\|$$

if a rank loss in downdating

$$bound(E) \leftarrow \sqrt{bound(E)^2 + \sigma_k^2}$$

Otherwise, no change. Similar change for update.

If $bound(E) > \epsilon$, compute $\|E\|$ by earlier method and see if refinement is possible. Reset $bound(E) = \|E\|$ in any case.

Numerical Tests

Created

$$
\begin{aligned}
X_{big} &\in \mathsf{R}^{150\times 20} \\
&= DY_{big}, \quad Y_{big}\, N(0,1) \\
D &= \mathrm{diag}(d_1, \ldots, d_{150}) \\
d_j &\in \{1, 10^{-3}, 10^{-5}, 10^{-7}, 10^{-9}\}
\end{aligned}
$$

30 rows each chosen at random for each weight.

At step k, maintained

$$
\begin{aligned}
X_{big}(k\colon k+29, :) &= Y\Sigma W^T, \qquad \text{SVD} \\
&= U_1 L V_1^T + U_2 \begin{pmatrix} F & G \end{pmatrix} V^T \qquad \text{ULVD} \\
&= U_1 L V_1^T + E \qquad \text{Truncated ULVD}
\end{aligned}
$$

Used update and downdate routines on last two. SVD was computed from scratch at each step. Choice of $\epsilon = 10^{-3}$.

Numerical Tests (2)

Looked at

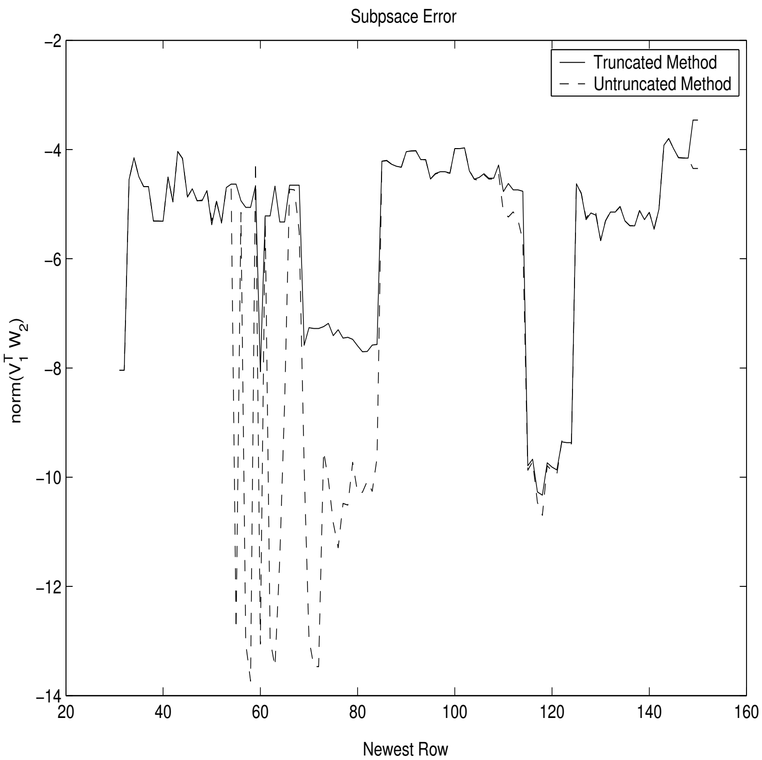
$$\|V_1^T W_2\|, \quad \text{subspace error}$$

$$\begin{array}{r} |bound(E) - \|E\|| \\ |bound(E) - \|\Sigma_2\|| \end{array}, \quad \text{norm errors}$$

$$\|U_1^T U_1 - I\|, \|V_1^T V_1 - I\|$$

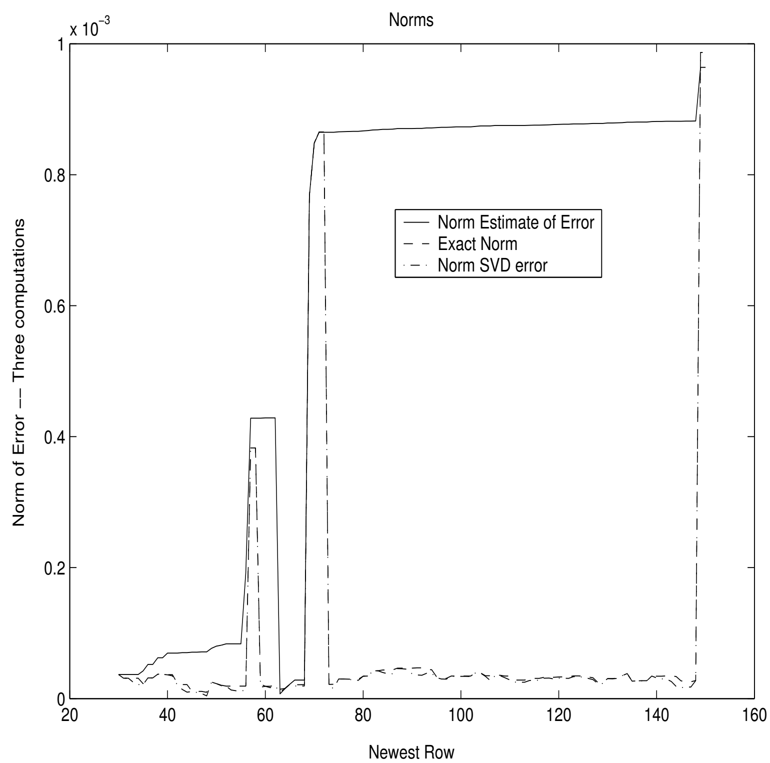
Orthogonal errors, Very boring.

Numerical Tests (3)– Subspace Errors



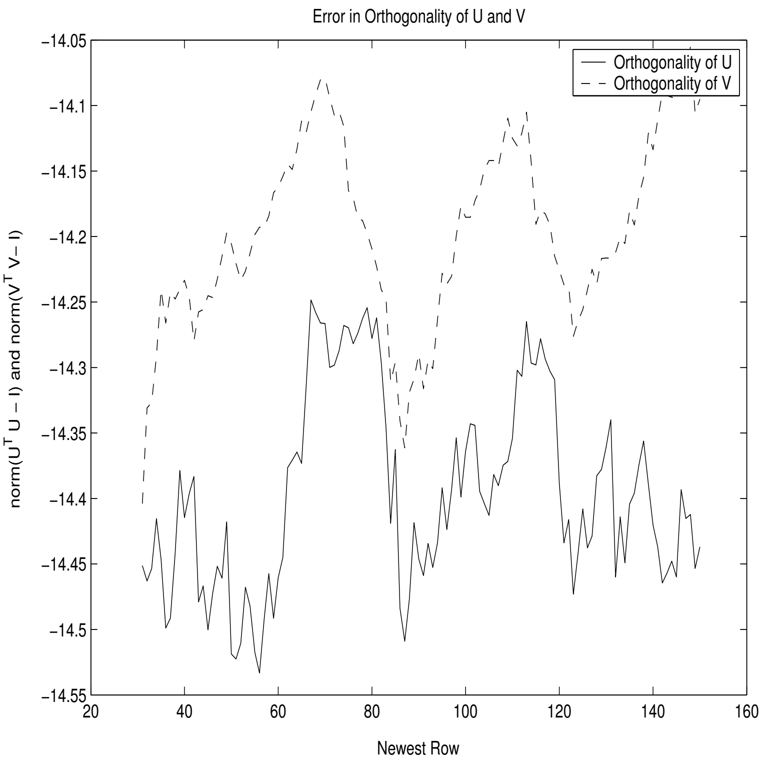
Errors in computing V_1

Numerical Tests (4)– Norm Estimate



Error in Norm Estimates

Numerical Tests (5)– Orthogonality of U_1 and V_1



Error in Orthogonality

Conclusions

It is both desirable and possible to maintain a matrix approximation of the form

$$X = U_1 L V_1^T + E$$

without explicitly storing E in any form.

- Update and downdate algorithms can formulated,
- Approximations can improved
- Rank can be estimated with reasonable dependability