The analytic singular value decomposition

Dáša Janovská
Vladimír Janovský

PANM 13, Prague 28.5. – 31.5.2006

Outline:

- **ASVD** concept
- **ASVD: Pathfollowing approach**
  Janovská, Janovský, Tanabe: ENUMATH 2005 (2006)
- **Singularities on the path**
  analysis

SVD ... *Singular Value Decomposition*

$A \in \mathbb{R}^{m \times n}$, $m \geq n$: $A = U \Sigma V^T$,

- $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ ... orthogonal matrices
- $\Sigma = \mathrm{diag}(s_1, \ldots, s_n) \in \mathbb{R}^{m \times n}$ ... diagonal matrix

s_i ... *singular values*,
$U_i \in \mathbb{R}^m$, $V_i \in \mathbb{R}^n$ ... *left/right singular vectors*

ASVD ... *Analytic Singular Value Decomposition*

= a version of SVD for
parameter dependent matrices

Kato, *Perturbation Theory for Linear Operators*, 1976.
Bunse-Gerstner, Byers, Mehrmann and Nichols: Numer.Math. 60 (1991)

Let $A = A(t)$, $A \in C^\omega([a,b], \mathbb{R}^{m \times n})$ ... real analytic

Construct a factorization

$$A(t) = U(t)\Sigma(t)V(t)^T$$

such that

- U, V and Σ are real analytic on $[a,b]$

For each $t \in [a,b]$:

- $U(t) \in \mathbb{R}^{m \times m}$, $V(t) \in \mathbb{R}^{n \times n}$ ... orthogonal matrices
- $\Sigma(t) = \text{diag}(s_1(t), \ldots, s_n(t)) \in \mathbb{R}^{m \times n}$ ... diagonal matrix

At $t = a$:

- $U(a)$, $\Sigma(a)$, $V(a)$ ... via SVD of $A(a)$

Specific properties:

The smoothness of $A(t) \Longrightarrow$ the singular values $s_i(t)$

- may be negative
- their ordering may by arbitrary

ASVD ... for large sparse matrices?

Apply

the pathfollowing of an implicitly defined curve

Janovská, Janovský, Tanabe: ENUMATH 2005 (2006)

- Consider *branches* of singular values $s_i(t) \in \mathbb{R}$ and corresponding left/right singular vectors $U_i(t) \in \mathbb{R}^m$, $V_i(t) \in \mathbb{R}^n$:

$$\begin{aligned} A(t)V_i(t) &= s_i(t)U_i(t)\,, \\ A(t)^T U_i(t) &= s_i(t)V_i(t)\,, \\ U_i(t)^T U_i(t) &= V_i(t)^T V_i(t) = 1 \end{aligned}$$

  for $t \in [a, b]$.

- Add the natural orthogonality conditions $U_i(t)^T U_j(t) = V_i(t)^T V_j(t) = 0$, $i \neq j$, $t \in [a, b]$.

- Consider just p, $p \leq n$, selected singular values $S(t) = (s_1(t), \ldots, s_p(t)) \in \mathbb{R}^p$, and the corresponding left/right singular vectors $U(t) = [U_1(t), \ldots, U_p(t)] \in \mathbb{R}^{m \times p}$, $V(t) = [V_1(t), \ldots, V_p(t)] \in \mathbb{R}^{n \times p}$ as smooth functions of $t \in [a, b]$.

In the operator setting,

$$F : \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p} \to \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p} \times \mathbb{R}^{p\times p} \times \mathbb{R}^{p\times p}$$

i.e.,

$$F : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^M , \quad N = p(1 + n + m) , \quad M = p(m + n + 2p) ,$$

where

$$\begin{aligned} F(t, X) &\equiv \Big(A(t)V - U\Sigma, A^T(t)U - V\Sigma, U^TU - I, V^TV - I\Big) , \\ X &\equiv (S, U, V) \in \mathbb{R}^p \times \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p} , \quad \Sigma \equiv \mathrm{diag}(S) \end{aligned}$$

$$F(t, X) = 0$$

... a system of *overdetermined nonlinear equations*

Computing the curve $F(t, X) = 0$

via a **predictor-corrector**

pathfollowing algorithm

In particular, *tangent continuation* is applied

Deuflhart, Hohmann,
Numerical Analysis in Modern Scientific Computing, 2003.

Experiment: homotopy

$$A(t) = t\,A2 + (1-t)\,A1\,, \quad t \in [0,1]$$

where

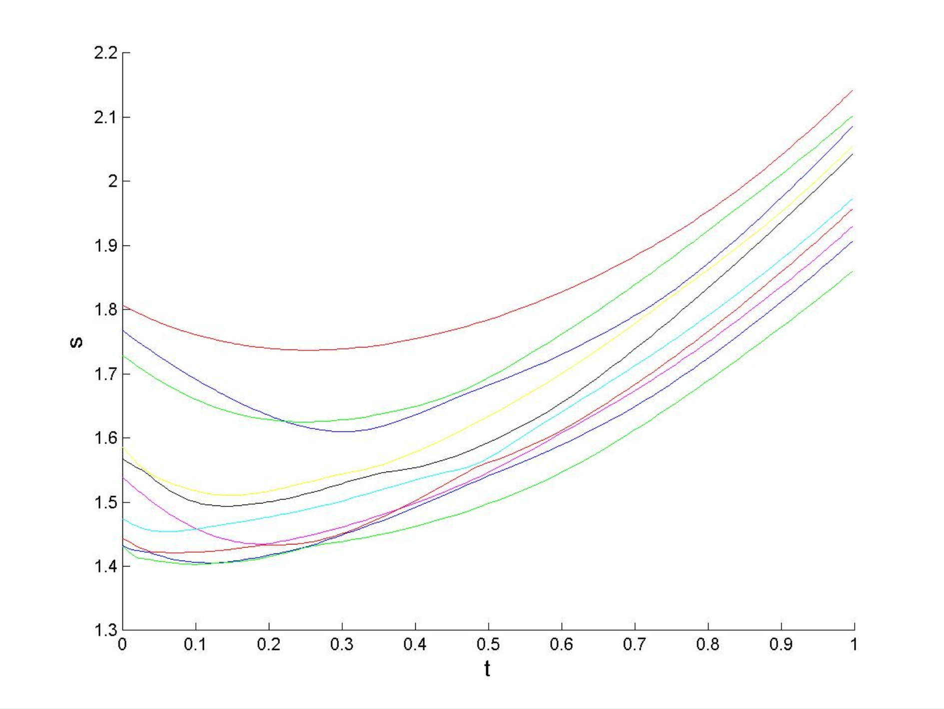
$$A1 \equiv \texttt{well1033.mtx}\,, \quad A2 \equiv \texttt{illc1033.mtx}$$

see ... `http://math.nist.gov/MatrixMarket/`

$A1, A2 \in \mathbb{R}^{1033\times 320}$, sparse, well/ill-conditioned

Initialization: at $t = 0$,
the 10 largest singular values, left/right singular vectors of $A1$
computed via `svds` ... see MATLAB Function Reference

Parameter t vs. the 10 largest singular values s:



Find path: $(t, X(t))$, $t \in [a, b]$,
$X(t) = (S(t), U(t), V(t)) \in \mathbb{R}^p \times \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p}$

Implementation of the algorithm

- Initialization: $(S(a), U(a), V(a))$ ... via **`svd`**, **`svds`**

  Assumption: $S(a) = (s_1(a), \ldots, s_p(a))$, $s_i(a)$ are simple for $i = 1, \ldots, p$.

  $\Longrightarrow$ a local existence of the branch

- **Singular points** on the path
  at such t: $s_i(t)$ ... nonsimple singular value

  In practice, the continuation may get stuck.

- **Remedies:** Extrapolation strategies
  - "Early Warning" of such values of t
  - "Jump Over" singular point

Recall:

Definition 1 *We say that* $s \in \mathbb{R}$ *is a* **singular value** *of the matrix* A *if there exist* $u \in \mathbb{R}^m$ *and* $v \in \mathbb{R}^n$ *such that*

$$Av - su = 0\,, \quad A^T u - sv = 0\,, \quad \|u\| = \|v\| = 1\,. \tag{1}$$

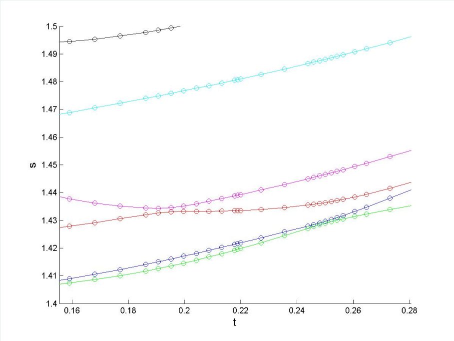
The vectors v *and* u *are called the right and the left singular vectors of the matrix* A*.*

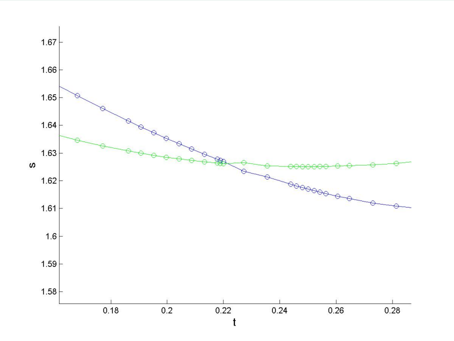
Definition 2 $s \in \mathbb{R}$ *is a* **simple singular value** *of the matrix* A *if there exist* $u \in \mathbb{R}^m$ *and* $v \in \mathbb{R}^n$ *such that*

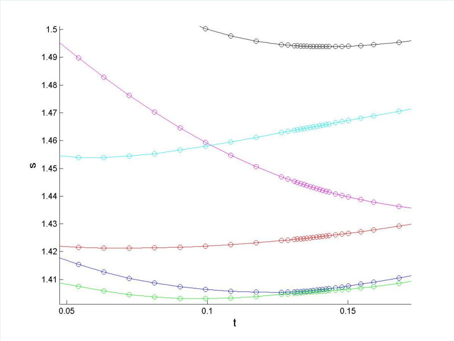
$$(s, u, v)\,, \quad (s, -u, -v)\,, \quad (-s, -u, v)\,, \quad (-s, u, -v)$$

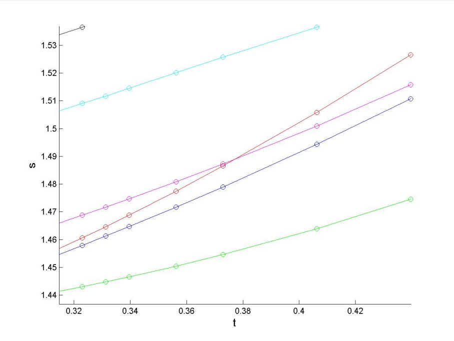
are, for the given s*, the only solutions to (1).*
A singular value s *which is not a simple singular value is called* **nonsimple singular value***.*

Zooms:









Branching scenario at singular points:

... the branches $t \longmapsto s_i(t)$ of singular values, $i = 1, \ldots, p$,
may intersect at isolated points only, namely,
at the points where

$$s_i(t) = s_j(t) \quad \text{or} \quad s_i(t) = -s_j(t)$$

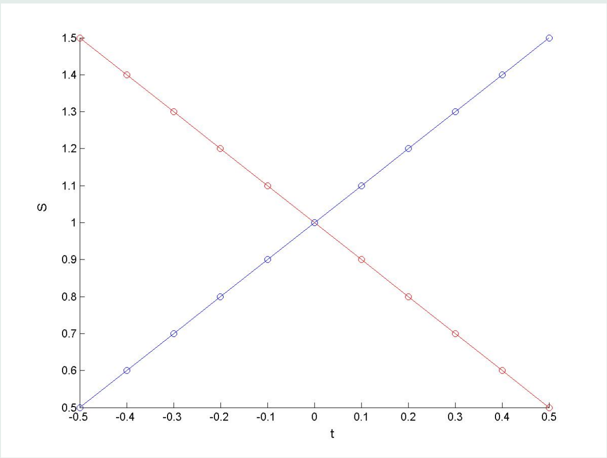
for $i \neq j$

Bunse-Gerstner, Byers, Mehrmann and Nichols: Numer.Math. 60 (1991)
Wright: Numer.Math. 60 (1992)

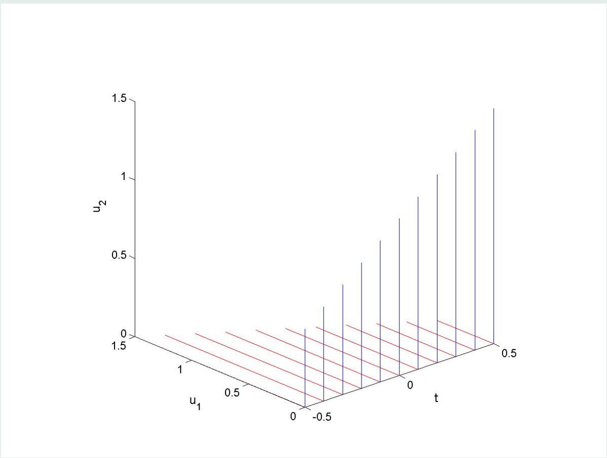
Experiment:

$$A(t) = \begin{pmatrix} 1-t & 0 \\ 0 & 1+t \end{pmatrix}$$

ASVD: $A(t) = U(t)\Sigma(t)V(t)^T$, $-0.5 \le t \le 0.5$.



$s_1(t), s_2(t)$



$s_1(t)\,U_1(t),\ s_2(t)\,U_2(t)$

at $t = 0$: $s_1(0)$, $s_2(0)$ ... nonsimple singular value of $A(0)$

The concept of genericity:

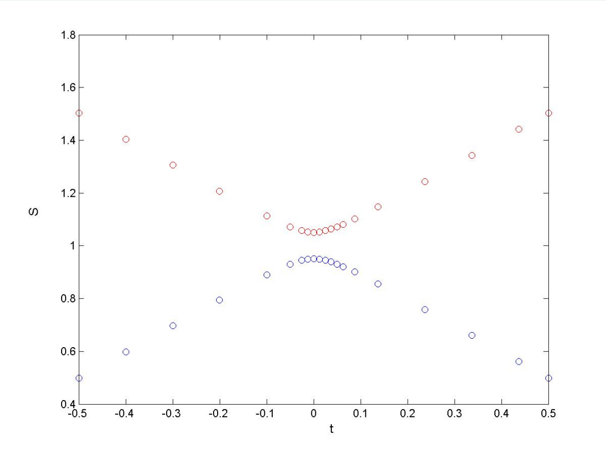
$$A^{\varepsilon}(t) \equiv A(t) + \varepsilon Z(t)\,, \quad a \le t \le b\,.$$

Does the ASVD-path persist
an arbitrary sufficiently small perturbation?

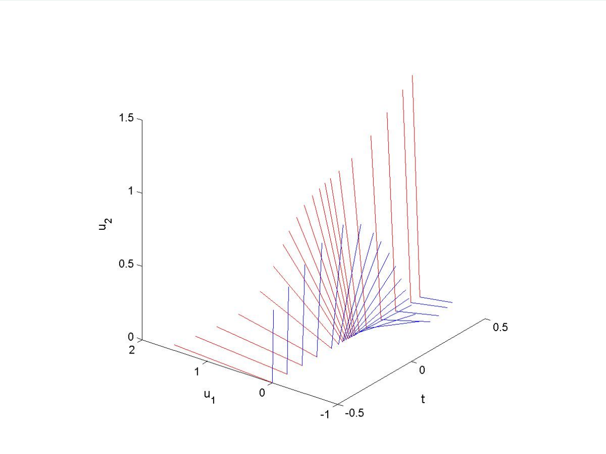
$$\begin{aligned} A(t) &= U(t)\Sigma(t)V(t)^T \\ A^{\varepsilon}(t) &= U^{\varepsilon}(t)\Sigma^{\varepsilon}(t)V^{\varepsilon}(t)^T \end{aligned}$$

Experiment continued:

$$A^\varepsilon(t) = \begin{pmatrix} 1-t & 0 \\ 0 & 1+t \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$



$s_1^\varepsilon(t), s_2^\varepsilon(t)$



$s_1^\varepsilon(t)\, U_1^\varepsilon(t),\ s_2^\varepsilon(t)\, U_2^\varepsilon(t)$

the case: $\varepsilon = 0.1$.

Reducing the problem

Recall: $F(t, X) = 0$, the overdetermined system
... branches of selected left/right singular vectors $U_i(t)$, $V_i(t)$
and corresponding singular values $s_i(t)$
plus orthogonality conditions

Note: Redundancy in scaling
either $U_i(t)^T U_i(t) = 1$ or $V_i(t)^T V_i(t) = 1$, or a combination,
become redundant
provided that $s_i(t) \neq 0$

Question: $\exists$ branch of
$s(t) \in \mathbb{R}$, $u(t) \in \mathbb{R}^m$, $v(t) \in \mathbb{R}^n$ ...
singular value, left/right singular vectors

Let

$$f : \mathbb{R} \times \mathbb{R}^{1+m+n} \rightarrow \mathbb{R}^{1+m+n}$$

$$t \in \mathbb{R}\,, \qquad x = (s, u, v) \in \mathbb{R}^1 \times \mathbb{R}^m \times \mathbb{R}^n \longmapsto f(t, x) \in \mathbb{R}^{1+m+n}$$

$$f(t, x) \equiv \begin{pmatrix} -su + A(t)v \\ A^T(t)u - sv \\ v^T v - 1 \end{pmatrix}$$

alternative:

$$f(t, x) \equiv \begin{pmatrix} -su + A(t)v \\ A^T(t)u - sv \\ u^T u + v^T v - 2 \end{pmatrix}$$

Problem 1 *Let* $t^0 \in \mathcal{J}$ *and*

$$x^0 = (s^0, u^0, v^0) : f(t^0, x^0) = 0\,.$$

Find all $(t, x) \in \mathbb{R} \times \mathbb{R}^{1+m+n}$: $f(t, x) = 0$.

... solution manifold

? ... a branch $f(t, x(t)) = 0$

Proposition 1 *Let* $s^0 \neq 0$ *be a simple singular value of* $A(t^0)$*.*
Then there exists a smooth

$$t \in \mathcal{I} \longmapsto x(t) \in \mathbb{R}^{1+m+n}$$

such that

$f(t, x(t)) = 0$ *for all* $t \in \mathcal{I}$

$x(t^0) = x^0$.

Treatments of singular points

Case: $s^0 \neq 0$ ... nonsimple singular points of $A(t^0)$
i.e.,
$f(t^0, x^0) = 0$, $x^0 = (s^0, u^0, v^0)$, $s^0 \neq 0$,
$\dim \mathrm{Ker}\, f_x(t^0, s^0) \geq 1$.
Let $\dim \mathrm{Ker}\, f_x(t^0, s^0) = 1$
$\Rightarrow$

$$\mathrm{Ker}\, f_x(t^0, x^0) = \mathrm{span} \left\{ \begin{pmatrix} 0 \\ \delta u \\ \delta v \end{pmatrix} \right\}, \; \|\delta u\| = \|\delta v\| = 1$$

Idea: apply **Singularity Theory**

Govaerts: Numerical Methods for Bifurcation ... , 2000

Ingredients: dimensional reduction
expanding bifurcation equation

The solution manifold $f(t, x) = 0$ is, locally, parametrizable via $\tau \in \mathbb{R}$, $\xi \in \mathbb{R}$ ... observable such that

$$f(t^0 + \tau, x^0 + \Delta x) = 0\,, \qquad \Delta x = \Delta x(\tau, \xi) \tag{2}$$

Moreover, (2) if and only if

$$\varphi(\tau, \xi) = 0 \tag{3}$$

The scalar equation (3) ... **bifurcation equation**

Analyzing bifurcation equation:
computing *leading terms* of Taylor expansion of $\varphi(\tau, \xi)$ at $(\tau, \xi) = 0 \in \mathbb{R}^2$

Theorem 1 ...

$\varphi_\tau = 0$, $\varphi_{\xi\tau} \neq 0$, $\varphi_{\tau\tau} \neq 0$ ... codim = 1 *singularity*
$\Rightarrow$ *branches*

$$\begin{aligned} \tau &= 0 \\ \xi &= -\frac{1}{2}\frac{\varphi_{\tau\tau}}{\varphi_{\xi\tau}}\tau + O(\tau^2) \end{aligned}$$

$O(\tau^2)$ *link to the solution manifold (2).*

The singular point under a perturbation: unfolding

e.g., let

$$A(t) + \varepsilon Z(t) \in \mathbb{R}^{m \times n}$$

? $f(t, x\,;\,\varepsilon) = 0$, $x = (s, u, v) \in \mathbb{R}^1 \times \mathbb{R}^m \times \mathbb{R}^n$,

$$f(t, x\,;\,\varepsilon) \equiv \begin{pmatrix} -su + (A(t) + \varepsilon Z(t))\, v \\ (A(t) + \varepsilon Z(t))^T\, u - sv \\ u^T u + v^T v - 2 \end{pmatrix}$$

Investigate corresponding bifurcation equation

$\varphi(\tau, \xi\,;\,\varepsilon) = 0\,, \quad (\tau, \xi) \in \mathbb{R}^2\,, \varepsilon \in \mathbb{R}$

Bifurcation equation: asymptotic analysis

$$\tau\left(\varphi_{\xi\tau}\,\xi+\frac{1}{2}\varphi_{\tau\tau}\,\tau+\text{h.o.t.}\right)+$$
$$+\,\varepsilon\,\left(\varphi_{\varepsilon}+\varphi_{\xi\varepsilon}\,\xi+\,\varphi_{\tau\varepsilon}\,\tau\,+\text{h.o.t.}\right)+$$
$$+\,O(\varepsilon^2)\,=\,0$$

where
$\varphi_\varepsilon = \frac{1}{2}\,\delta u^T\,Z(t^0)\,v^0\,+\,\frac{1}{2}\,(u^0)^T\,Z(t^0)\,\delta v^T$
etc.

Conclusions

- **ASVD** via a pathfollowing
- Algorithm may get stuck at isolated points $\equiv$ singular points
- Singular points can be
  - investigated by reducing the problem to **bifurcation equation**
  - approximated by computing leading terms of bifurcation equation
- Ad singular points: Instead of Extrapolation strategies, propose a numerical treatment based on Singularity Theory