

Conditions of Matrices in Discrete Tension Spline Approximations of DMBVP

Mladen Rogina, Sanja Singer

October 9, 2006

Abstract

Some splines can be defined as solutions of differential multi-point boundary value problems (DMBVP). In the numerical treatment of DMBVP, the differential operator is discretized by finite differences. We consider one dimensional discrete hyperbolic tension spline introduced in [2], and the associated specially structured pentadiagonal linear system.

Error in direct methods for the solution of this linear system depends on condition numbers of corresponding matrices. If the chosen mesh is uniform, the system matrix is symmetric and positive definite, and it is easy to compute both, lower and upper bound, for its condition. In the more interesting non-uniform case, matrix is not symmetric, but in some circumstances we can nevertheless find an upper bound on its condition number.

1 Introduction

In [2] Costantini *et al.* introduced discrete hyperbolic tension splines as a generalization of discrete cubic splines, which were mentioned for the first time by Malcolm [8]. The idea of *univariate discrete tension spline* is the following: Given $n_i \in \mathbb{N}$, $i = 0, \dots, N$, find a discrete function u_{ij} , $j = -1, \dots, n_i + 1$, $i = 0, \dots, N$ satisfying the difference equations:

$$\left[\Lambda_i^2 - \left(\frac{p_i}{h_i} \right)^2 \Lambda_i \right] u_{ij} = 0, \quad j = 1, \dots, n_i - 1, \quad i = 0, \dots, N, \quad (1)$$

where

$$\Lambda_i u_{ij} = \frac{u_{i,j-i} - 2u_{i,j} + u_{i,j+1}}{\tau_i^2}, \quad \tau_i = \frac{h_i}{n_i},$$

subject to the *discrete smoothness conditions*:

$$\begin{aligned} u_{i-1, n_{i-1}} &= u_{i,0} \\ \frac{u_{i-1, n_{i-1}+1} - u_{i-1, n_{i-1}-1}}{2\tau_{i-1}} &= \frac{u_{i,1} - u_{i,-i}}{2\tau_i}, \quad i = 1, \dots, N \\ \Lambda_{i-1} u_{i-1, n_{i-1}} &= \Lambda_i u_{i,0} \end{aligned}$$

where $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$ are the biggest and the smallest singular value of A . If A is symmetric and positive definite (10) is equivalent to

$$\kappa_2(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)},$$

where $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ are the biggest and the smallest eigenvalue of A . Furthermore, we can compare eigenvalues of matrices by the Weyl's theorem [4], pp. 181:

Theorem 1 *Let $A, B \in C^{n \times n}$ be Hermitian and let the eigenvalues $\lambda_i(A)$, $\lambda_i(B)$ and $\lambda_i(A + B)$ be arranged in increasing (in fact nondecreasing) order*

$$\lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = \lambda_{\max}.$$

For each $k = 1, \dots, n$ we have

$$\lambda_k(A) + \lambda_1(B) \leq \lambda_k(A + B) \leq \lambda_k(A) + \lambda_n(B).$$

For our purposes, let us substitute $k = 1$, $A = C$, $B = D$ in Weyl's theorem to obtain

$$\lambda_{\min}(C) + \lambda_{\min}(D) \leq \lambda_{\min}(A) \leq \lambda_{\min}(C) + \lambda_{\max}(D), \quad (11)$$

and $k = 1$, $A = D$, $B = C$ to obtain

$$\lambda_{\min}(C) + \lambda_{\min}(D) \leq \lambda_{\min}(A) \leq \lambda_{\min}(D) + \lambda_{\max}(C). \quad (12)$$

By substituting $k = n$, $A = C$, $B = D$ we have

$$\lambda_{\max}(C) + \lambda_{\min}(D) \leq \lambda_{\max}(A) \leq \lambda_{\max}(C) + \lambda_{\max}(D), \quad (13)$$

and finally, substitution of $k = n$, $A = D$, $B = C$ gives

$$\lambda_{\min}(C) + \lambda_{\max}(D) \leq \lambda_{\max}(A) \leq \lambda_{\max}(C) + \lambda_{\max}(D). \quad (14)$$

Relations (11) and (12) give

$$\begin{aligned} \lambda_{\min}(C) + \lambda_{\min}(D) &\leq \lambda_{\min}(A) \\ &\leq \min\{\lambda_{\min}(C) + \lambda_{\max}(D), \lambda_{\min}(D) + \lambda_{\max}(C)\}, \end{aligned} \quad (15)$$

and, similarly, (13) and (14) give

$$\begin{aligned} \max\{\lambda_{\max}(C) + \lambda_{\min}(D), \lambda_{\min}(C) + \lambda_{\max}(D)\} &\leq \lambda_{\max}(A) \\ &\leq \lambda_{\max}(C) + \lambda_{\max}(D). \end{aligned} \quad (16)$$

From the structure of C in (8)–(9), we obtain

$$\begin{aligned} \lambda_{\min}(C) &= \min_i \left[4 \left(1 - \cos \frac{\pi}{n_i} \right)^2 + 2\omega_i \left(1 - \cos \frac{\pi}{n_i} \right) \right], \\ \lambda_{\max}(C) &= \max_i \left[4 \left(1 + \cos \frac{\pi}{n_i} \right)^2 + 2\omega_i \left(1 + \cos \frac{\pi}{n_i} \right) \right], \\ \lambda_{\min}(D) &= 0, \\ \lambda_{\max}(D) &= 2. \end{aligned}$$

By substituting these eigenvalues into (15) and (16) we obtain

$$\begin{aligned} \lambda_{\min}(C) &\leq \lambda_{\min}(A) \leq \min\{\lambda_{\min}(C) + 2, \lambda_{\max}(C)\}, \\ \max\{\lambda_{\max}(C), \lambda_{\min}(C) + 2\} &\leq \lambda_{\max}(A) \leq \lambda_{\max}(C) + 2, \end{aligned}$$

and

$$\frac{\max\{\lambda_{\max}(C), \lambda_{\min}(C) + 2\}}{\min\{\lambda_{\min}(C) + 2, \lambda_{\max}(C)\}} \leq \kappa_2(A) \leq \frac{\lambda_{\max}(C) + 2}{\lambda_{\min}(C)}. \quad (17)$$

In addition, simple upper bound for the $\lambda_{\max}(A)$ can be obtained by Gershgorin's theorem

$$\lambda_{\max}(A) \leq 16 + 4\omega_i. \quad (18)$$

Coupling (17) and (18) together, we obtain

$$\frac{\max\{\lambda_{\max}(C), \lambda_{\min}(C) + 2\}}{\min\{\lambda_{\min}(C) + 2, \lambda_{\max}(C)\}} \leq \kappa_2(A) \leq \frac{\min\{\lambda_{\max}(C) + 2, 16 + 4\omega_i\}}{\lambda_{\min}(C)}.$$

We have estimated conditions of matrices A with various relationships between p_i and n_i . As the reference point we calculated spectral condition number of each A by using accurate SVD [3].

Example 1 *Let us take test matrices A all of order 60, with equal tension parameters p_i for all blocks, but with different structures of inner blocks. The first family of matrices consists of matrices with ten blocks C_i of order 6, while in the second family the block C_0 is of order 24, and the other nine blocks are of order 4. Our estimator gives:*

p_i	0.0	0.01	0.1	1.0	10	100
$\kappa_2(A)$	375.0	375.0	374.6	341.7	50.6	19.5
lower bound	7.1	7.1	7.1	7.1	9.1	18.6
upper bound	407.9	407.9	407.5	371.7	54.5	19.6

Table 1: Estimates for the first family of matrices A .

p_i	0.0	0.01	0.1	1.0	10	100
$\kappa_2(A)$	31713.9	31717.7	31694.6	29905.2	8664.1	5756.6
lower bound	7.9	7.9	7.9	7.9	13.8	648.3
upper bound	64331.6	64331.6	64272.7	58990.0	10664.6	5789.5

Table 2: Estimates for the second family of matrices A .

As expected, conditions of matrices with equal-sized blocks are lower than conditions of matrices with blocks of widely varying size. Moreover, for sufficiently large p_i we are very close to the reference condition number.

Nonsymmetric case

For non-uniform meshes, we proceed in the same way, by considering the splitting $A = C + E$, where symmetric, positive definite C is equal to C from the symmetric case, and E is the nonsymmetric replacement for D . We have

$$\|A\|_2 = \|C + E\|_2 \leq \|C\|_2 + \|E\|_2 = \lambda_{\max}(C) + \sqrt{\lambda_{\max}(E^T E)}. \quad (19)$$

Diagonal blocks of E are

$$\begin{bmatrix} \frac{2}{\rho_i + 1} & \frac{2}{\rho_i(\rho_i + 1)} \\ \frac{2\rho_i^2}{\rho_i + 1} & \frac{2\rho_i}{\rho_i + 1} \end{bmatrix} \quad (20)$$

instead of $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ in D . It is easy to compute that

$$(E^T E)_i = \frac{4(1 + \rho_i^4)}{(\rho_i + 1)^2} \begin{bmatrix} 1 & \frac{1}{\rho_i} \\ \frac{1}{\rho_i} & \frac{1}{\rho_i^2} \end{bmatrix}.$$

Also, if λ is eigenvalue of A , then $k\lambda$ is eigenvalue of kA , and we need to compute the eigenvalues of

$$\begin{bmatrix} 1 & \frac{1}{\rho_i} \\ \frac{1}{\rho_i} & \frac{1}{\rho_i^2} \end{bmatrix},$$

which are readily found to be

$$\lambda_{\min} = 0, \quad \lambda_{\max} = \frac{1 + \rho_i^2}{\rho_i^2},$$

thus

$$\lambda_{\max}(E^T E) = \max_i \left(\frac{4(1 + \rho_i^4)}{(\rho_i + 1)^2} \cdot \frac{1 + \rho_i^2}{\rho_i^2} \right) = \max_i \frac{4(1 + \rho_i^4)(1 + \rho_i^2)}{\rho_i^2(\rho_i + 1)^2}.$$

Previous formula, together with (19) gives

$$\|A\|_2 \leq \lambda_{\max}(C) + \max_i \frac{2}{\rho_i(\rho_i + 1)} \sqrt{(1 + \rho_i^4)(1 + \rho_i^2)}.$$

We also need to bound $\|A^{-1}\|_2$. According to Corollary 3.1.5. from [5], if singular values of A and eigenvalues of $H(A) = \frac{1}{2}(A + A^*)$ are nonincreasingly ordered, for each singular value of A we have

$$\sigma_k(A) \geq \lambda_k(H(A)), \quad k = 1, \dots, n. \quad (21)$$

On the other hand we can write $H(A)$ as

$$H(A) = \frac{1}{2}(A + A^T) = C + \frac{1}{2}(E + E^T).$$

Let $F = \frac{1}{2}(E + E^T)$. By Weyl's theorem, we obtain lower bound for $H(A)$:

$$\lambda_{\min}(C) + \lambda_{\min}(F) \leq \lambda_{\min}(H(A)). \quad (22)$$

Now (10), (19) and (22) yield

$$\kappa_2(A) \leq \frac{\|A\|_2}{\lambda_{\min}H(A)} \leq \frac{\lambda_{\max}(C) + \sqrt{\lambda_{\max}(E^T E)}}{\lambda_{\min}(C) + \lambda_{\min}(F)}. \quad (23)$$

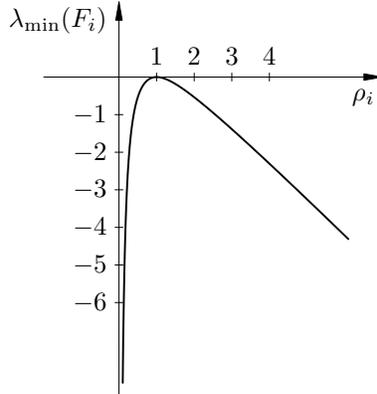
It remains to derive a lower bound for $\lambda_{\min}(F)$. From (20), it is easy to calculate that diagonal blocks F_i of F have the form

$$\begin{bmatrix} \frac{2}{\rho_i + 1} & \frac{1 + \rho_i^3}{\rho_i(\rho_i + 1)} \\ \frac{1 + \rho_i^3}{\rho_i(\rho_i + 1)} & \frac{2\rho_i}{\rho_i + 1} \end{bmatrix}.$$

Eigenvalues of F are zeros and

$$\lambda(F_i) = 1 \pm \frac{\sqrt{(\rho_i^4 + 1)(\rho_i^2 + 1)}}{\rho_i(\rho_i + 1)}.$$

Also, it is easy to check that $\min \lambda(F_i) \leq 0$ and $\min \lambda(F_i) = 0$ if and only if $\rho_i = 1$. For each ρ_i we have the following graph:



Obviously, minimal $\lambda_{\min}(C) + \lambda_{\min}(F)$ is non-negative in some small neighbourhood of 1 depending on n_k and ω_k .

Example 2 *There exist nonsymmetric matrices such that (23) is useless (denominator of the right-hand side is less than 0), their condition being much higher than the condition of corresponding symmetric matrices.*

For example, if A has seven blocks of order 6, with p_i 's equal to 1 for all blocks, and $h = (4.2, 0.1, 4.1, 0.5, 4.1, 0.1, 4.4)$, then $\kappa_2(A) = 59083.4$. If we change h such that $h_i = n_i$ for all i , i.e., if A is symmetric, then the condition of a such matrix is 341.6.

If we have sufficiently big p_i 's, and if put $n_i \approx h_i$, then instead of $\tau = \tau_i = 1$ we have $\tau_i \approx 1$, and conditions do not differ significantly from conditions of corresponding symmetric matrices.

For example, we can take matrix A with 6 blocks of order 4, and $p_i = 12$ for all i . For symmetric A_s we could take, for example, $h_i = 5$ for all i . The following table gives conditions and their upper bounds for both symmetric A_s , and slightly perturbed nonsymmetric A_n with $h = (5.01, 4.9, 5.01, 5.0, 4.95, 5.2)$.

$\kappa_2(A_s)$	bound for $\kappa_2(A_s)$	$\kappa_2(A_n)$	bound for $\kappa_2(A_n)$
14.6976	15.3153	14.6985	15.3355

Table 3: Conditions of symmetric A_s and “close” nonsymmetric A_n .

3 Componentwise perturbations

For a linear system $Ax = b$, perturbations ΔA and Δb such that $(A + \Delta A)(x + \Delta x) = b + \Delta b$ are componentwise perturbations if

$$|\Delta A| \leq \varepsilon |A|, \quad |\Delta b| \leq \varepsilon |b|,$$

where $|\cdot|$ denotes pointwise absolute value ($|A|_{ij} = |A_{ij}|$) for some $\varepsilon > 0$. Note that componentwise perturbations will not perturb zeros in A and b .

Skeel ([11], Theorem 2.1) shows that

$$\frac{\|\Delta x\|_\infty}{\|x\|_\infty} \leq \varepsilon \frac{\| |A^{-1}| |A| |x| + |A^{-1}| |b| \|_\infty}{(1 - \varepsilon \| |A^{-1}| |A| \|_\infty) \|x\|_\infty},$$

introducing

$$\text{cond}(A, x) := \frac{\| |A^{-1}| |A| |x| \|_\infty}{\|x\|_\infty},$$

and an upper bound for $\text{cond}(A, x)$ as

$$\text{cond}(A) = \| |A^{-1}| |A| \|_\infty.$$

If D is the row scaling of A such that DA has unit 1-norm, Chandrasekaran and Ipsen in [1] note that

$$\frac{\kappa_\infty(A)}{\kappa_\infty(D)} \leq \text{cond}(A) \leq \kappa_\infty(A).$$

This shows that $\text{cond}(A) \approx \kappa_\infty(A)$ if rows of A are not badly scaled.

If A of order n is symmetric and positive definite, it is easy to bound $\text{cond}(A)$ by using eigenvalue decomposition of A , $A = U\Lambda U^T$, where U is unitary and Λ is diagonal matrix of eigenvalues. Then we have

$$|A^{-1}| = |U\Lambda^{-1}U^T| \leq |U||\Lambda^{-1}||U^T|,$$

It is easy to show that

$$|A^{-1}|_{ij} \leq \sum_{k=1}^n \frac{1}{\lambda_k(A)} |u_{ik}u_{jk}| \leq \sum_{k=1}^n \frac{1}{\lambda_k(A)} =: \mu.$$

Now,

$$|A^{-1}| \leq \mu G, \quad G = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix},$$

and

$$(|A^{-1}||A|)_{ij} \leq \mu G|A| = \mu \sum_{i=1}^n |A|_{ij}. \quad (24)$$

For fixed j , the right-hand side of (24) does not depend on i , and therefore

$$\text{cond}(A) \leq \mu \sum_{j=1}^n \sum_{i=1}^n |A|_{ij}. \quad (25)$$

If A is symmetric and defined by (2), then $n = \sum_{i=0}^N (n_i - 1)$. From (9) and Weil's theorem it follows that

$$\mu \leq \sum_{i=0}^N \sum_{j=1}^{n_i-1} \frac{1}{\lambda_j(C_i)}.$$

From (25) and (2)–(7) we obtain

$$\begin{aligned}
\text{cond}(A) &\leq \mu \left[\sum_{i=0}^N [(n_i - 3)b_i + 2(n_i - 2)|a_i| + 2(n_i - 2)] + (b_0 - 1) + (b_N - 1) \right. \\
&\quad \left. + \sum_{i=1}^N \eta_{i-1, n_{i-1}-1} + \sum_{i=1}^N \eta_{i,1} + \sum_{i=1}^N \delta_{i-1, n_{i-1}-1} + \sum_{i=1}^N \delta_{i,1} \right] \\
&= \mu \left[\sum_{i=0}^N (16n_i - 40 + 4n_i\omega_i - 10\omega_i) \right. \\
&\quad \left. + 10 + 2\omega_0 + 2\omega_N + \sum_{i=1}^N (6 + 2\omega_{i-1}) + \sum_{i=1}^N (6 + 2\omega_i) + 2N \right] \\
&= \mu \left[\sum_{i=0}^N (16n_i - 40 + 4n_i\omega_i - 10\omega_i) + 10 + 4 \sum_{i=0}^N \omega_i + 14N \right] \\
&= \mu \left[16 \sum_{i=0}^N n_i + 4 \sum_{i=0}^N n_i\omega_i - 6 \sum_{i=0}^N \omega_i - 26N - 30 \right].
\end{aligned}$$

Comparing bounds $\kappa_\infty(A) \leq \sqrt{n} \cdot \kappa_2(A)$ and $\text{cond}(A)$ may not be easy. Also, if A is nonsymmetric, no similar techniques exist to obtain componentwise bounds.

4 Conclusion

It is not always true that discretized DMBVP is well conditioned; it depends on n_i and ω_i . The ill-conditioning appearing for widely varying block sizes reflects the ill-posedness of the interpolation problem in which data points are dense in one region, and sparse in another. We have tested various cases and estimated the condition number using accurate SVD [3]. Numerical experiments seem to be in accordance with the *a priori* estimates we have obtained.

Since the choice of tension parameters comes from practical applications, like shape preserving approximation (see [6] and references therein), it is our hope that such a choice of tension parameters can be made, that both, shape-preserving requirements, and numerical stability can be achieved. The delicate balance between the two is at this moment not completely understood.

Acknowledgments This research was supported by Grant 0037114, by the Ministry of Science, Education and Sports of the Republic of Croatia.

References

- [1] S. Chandrasekaran and I. C. F. Ipsen, *On the Sensitivity of Solution Components in Linear Systems of Equations*, SIAM J. Matrix Anal. Appl. 16 (1), 1995, 93–112.

- [2] P. Costantini, B. I. Kvasov, C. Manni, *On Discrete Hyperbolic Tension Splines*, Adv. Comp. Math. 11, 1999, 331–354.
- [3] Z. Drmač, *Computing the Singular and the Generalized Singular Values*, Ph.D. Thesis, FernUniversität–Gesamthochschule, Hagen, 1994.
- [4] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1992.
- [5] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [6] B. I. Kvasov, *Methods of Shape-Preserving Spline Approximation*, World Scientific, Singapore, 2000.
- [7] B. I. Kvasov, *DMBVP for Tension Splines*, in *Proceedings of the Conference on Applied Mathematics and Scientific Computing*, June 23–27, 2003, M. Marušić, Z. Drmač and Z. Tutek (eds.), Springer, Dordrecht, 2005, 67–94.
- [8] M. A. Malcolm, *On the Computation of Nonlinear Spline Functions*, SIAM J. Numer. Anal. 14 (2), 1977, 255–281.
- [9] B. J. McCartin, *Theory of Exponential Splines*, J. Approx. Theory 66 (23), 1999, 1–23.
- [10] P. Rentrop, *An Algorithm for the Computation of the Exponential Splines*, Numer. Math. 35, 1980, 81–93.
- [11] R. D. Skeel, *Scaling for Numerical Stability in Gaussian Elimination*, J. Assoc. Comput. Mach. 26 (3), 1979, 494–526.
- [12] G. A. Watson, *Approximation Theory and Numerical Methods*, Wiley, 1980.