Extremely accurate solutions using block decomposition and extended precision for solving very ill-conditioned equations

†Kansa, E.J.[1], Skala V.[2], and Holoborodko, P.[3]

[1]Convergent Solutions, Livermore, CA 94550 USA
[2]Computer Science & Engineering Dept., Faculty of Applied Sciences, University of West Bohemia, University 8, CZ 301 00 Plzen, Czech Republic
[3]Advanpix LLC, Maison Takashima Bldg. 2F, Daimachi 10-15-201, Yokohama, Japan 221-0834
†Corresponding author: edwardjkansa@gmail.com

Abstract

Many authors have complained about the ill-conditioning associated the numerical solution of partial differential equations (PDEs) and integral equations (IEs) using as the continuously differentiable Gaussian and multiquadric continuously differentiable (C^∞) radial basis functions (RBFs). Unlike finite elements, finite difference, or finite volume methods that lave compact local support that give rise to sparse equations, the C^∞ -RBFs with simple collocation methods give rise to full, asymmetric systems of equations. Since C^∞ RBFs have adjustable constent or variable shape parameters, the resulting systems of equations that are solve on single or double precision computers can suffer from "ill-conditioning". Condition numbers can be either the absolute or relative condition number, but in the context of linear equations, the absolute condition number will be understood. Results will be presented that demonstrates the combination of Block Gaussian elimination, arbitrary arithmetic precision, and iterative refinement can give remarkably accurate numerical salutations to large Hilbert and van der Monde equation systems.

1. Introduction

An accurate definition of the condition number, κ, of the matrix, **A**, is the ratio of the largest to smallest absolute value of the singular values, $\{\sigma_i\}$, obtained from the singular value decomposition (SVD) method, see [1]:

$$\kappa(\mathbf{A}) = \max_j |\sigma_j| / \min_j |\sigma_j| \quad (1)$$

This definition of condition number will be used in this study.

Whenever the absolute condition number is comparable to the inverse of the machine epsilon, see [2, 3,4], the numerical results become unreliable. In some instances, the absolute condition number may be exceptionally large, but the relative is small.

The important cause of ill-conditioning is the number of bits assigned to a computer word. The Institute of Electrical and Electronics Engineers (IEEE) defined a single precision word to have 8 bits and double precision word to have 16 bits, and (the double precision maximum condition number is O(1e16). Since scientific computing comprises a minuscule fraction of the user market, extended memory chips are unlikely, hence software methods are needed to obtain extended precision.

C^∞ RBFs have the advantage of being deponently convergent and converges faster as the dimensions increases. Socially important problems such as controlled fusion [3], designing new medical drugs, option markets, etc. requires numerical solutions of higher dimensional

PDEs and IEs made more difficult because operator splitting is not a viable approach. Because of the curse of dimensionality, a f the fewer the fewer the number of points per dimension the better.
The approach that we are advocating is using C^{∞} RBFs, shape parameters that make the RBFs flatter, minimizing the number of data centers, using extended arithmetic precision and employing block decomposition to obtain many smaller ranked block matrices, each of which is better conditioned
Assume the original matrix of rank N is subdivided into K blocks, each of which contains P points. Thus, the matrix is subdivided as:

$\mathbf{A}_{1,1}$	$\mathbf{A}_{1,2}$	$\mathbf{A}_{1,3}$	$\mathbf{A}_{1,4}$	…	$\mathbf{A}_{1,K}$
$\mathbf{A}_{2,1}$	$\mathbf{A}_{2,2}$	$\mathbf{A}_{2,3}$	$\mathbf{A}_{2,4}$	…	…
$\mathbf{A}_{3,1}$	$\mathbf{A}_{3,2}$	$\mathbf{A}_{3,3}$	$\mathbf{A}_{3,4}$	…	…
…	…	…	…	…	…
$\mathbf{A}_{K,1}$	$\mathbf{A}_{K,2}$	$\mathbf{A}_{K,3}$	$\mathbf{A}_{K,4}$	…	$\mathbf{A}_{K,K}$

Figure 1. The full matrix A is partitioned to K blocks each of which contains P points.

Each block, $\mathbf{A}_{j,k,}$ is a square k×k, matrix that is a submatrix of the original N × N matrix, **A**.

There are various possible block decomposition methods available: block Gaussian decomposition, block singular value (SVD) decomposition, and block quotient- is a remainder (QR) decomposition. The simplest block decomposition method is the block Gaussian elimination method (BGEM) analog without pivoting. The BGEM is combined with extended arithmetic precision, and iterative refinement. The resulting block operations transform the original fully populated block matrix into identity block diagonal matrices and zero block matrices on the off-diagonal matrices. In additions, all operations are vectorized for maximum computational efficiency.

2. Example Test Problem of ill-conditioned linear equations

The test problem is the notoriously ill-conditioned Hilbert matrix whose elements are:

$$\mathbf{H}(i,j) = 1/(i + j - 1), \tag{2}$$

where i is i-th row index and j is the column index. The Hilbert matrix is invertible to all orders, and the inverse Hilbert matrix exists. However, on a double precision computer, the condition number is O(1e16). A heuristic rule for ill-conditioning is that the condition number increase with rank of the matrix.
If the unknown vector, **x**. is specified as $\mathbf{x} = [1,1\ldots,1]^{T}$, the right vector, **b**, is found by multiplying

$$\mathbf{b} = \mathbf{H}\mathbf{x}. \tag{3}$$

The problem that is considered is to assume the right-hand-side, **b**, is known, and to find x

A more accurate definition of the condition number of the matrix, **A**, is the ratio of the largest to smallest absolute value of the singular values, $\{\sigma i\}$, obtained from the singular value decomposition (SVD) method:

$$\kappa(A) = \max_j |\sigma_j| / \min_j |\sigma_j,| \quad (4)$$

This definition of condition number will be used.

3. Example calculations

The first set of numerical experiments examines the root mean square errors RMS)

$$\text{RMS error} = [\ \Sigma_i (x^{\text{exact}}_j - x^{\text{numerical}}_j)^2 / N\ ]^{1/2} \quad (5)$$

The first set of examples is the block partitioning a 300×300 Hilbert matrix with 48 and 200 digits of extended precision with various sized blocks.
Table 1. Comparison of RMS errors of a 300×300 Hilbert matrix using different block sizes digits of precision.

Table 1. RMS errors for block decomposition of a (300×300) Hilbert matrix with 48 and 200 digits of precision.

48 digits	Block size	RMS error	200 digits	Block size	RMS error
	10 (30×30)	0.73		10 (30×30)	7.1e-527
	15(20×20)	2,72e-15		15(20×20)	3.9e-527
	20(15×15)	2.19e-16		20(15×15)	4.3e-527
	30(10×10)	1.95e-20		30(10×10)	6.5e-537
	60 (5×5)	1.26e-24		60 (5×5)	4.9e-537

One last example was a 1000 × 1000 Hilbert matrix, with the right-hand vector being generated by the solution, $x= [1,1\ldots,1]^T$ in which 50 (20×20) blocks in which 200 digits of precision were used. Such a problem required so much memory that the run time required over 42hrs. The RMS error was 2.72e-126.
Because of the extremely slow execution speed observed with the 1000×1000 Hilbert matrix, no additional tests were conducted.
I some examples of the 300×300 Hilbert matrix, partitioning the original matrix into smaller blocks. The full the 300×300 Hilbert matrix, condition number is 1.1e20. For the following block sizes, the maximum block condition numbers are:

Table 2: Block size and maximin condition number of a 300×300 Hilbert matrix

Block size	Maximum condition number
(20 ×20)	2.23e+18
((15 × 15)	2.59e+17
(10 × 10)	1.60e+13

An alternative method to the block Gaussian elimination method involving P blocks containing a uniform number of k elements per block, it the decomposition of a larger matrix into many 2× 2 block matrices, see [10].
The next examples are the van der Monde matrices defined as a matrix of vectors raised to a power, commonly given by:

$$\mathbf{A(i,j) = v(j)^{(N-j)}} \tag{6}$$

We considered the vector composed of N elements starting with 1 and increasing by increments of 0.5. Many other variations can be considered.
We considered a vector of 100 elements ranging from 1.0 to 50.5, incremented by 0.5. The van der Monde matrix of rank 100 of such a vector is an estimated condition number of 8e+202. With 360 digits of precision, we considered 25 (4x4) blocks with a RMS error of 4e-44, and with 50 (2x2) blocks, the RMS error is 4e-95. So even with a horribly conditioned van der Monde matrix, we are still able to obtain very accurate RMS errors.

4. Discussion

Traditional domain decomposition methods are intrinsically iterative in nature whether they are overlapping or non-overlapping methods. A large domain over a large PDE or IE problems is subdivided into many smaller sub-domains. Artificial boundaries are constructed on which artificial boundary conditions are imposed, and the solution is obtained iteratively. For elliptic or time dependent problems for diffusion or viscosity dominates, the iterative matching of the function and normal and tangential derivatives can achieve a sufficient degree of convergence. The basic problem is that there is not a sufficient number of equations for the number of unknowns to enforce all the derivative continuity conditions.

For advective dominated problems, convergence is more of a problem unless one is willing to sacrifice physics to additional numerical viscosity by way of up wind spatial differencing. Domain decomposition methods embody large potential for a parallelization of the finite difference, finite element, or finite volume methods for distributed, parallel computations.

A non-iterative domain decomposition method was introduced in which large submatrices were solved individually and potentially in parallel with linear equation matching at surfaces, lines, and points, see]11]. Because of the different sizes of each type of sub-domain, parallelization is hampered.

Additional acceleration can be obtained by classical iterative refinement, and the geometric projective algebra, see [12].

These examples illustrate the well-known fact that the condition number depends upon the rank of the system of equations, and the number of digits available. However, as the numbers in the set of equations approaches the ideal Platonic limit of infinity, the time required also approaches infinity. The computation of important applications such as plasma fusion, designer medicine based upon the first principles of quantum mechanics, etc. will involve multi-dimensional calculations and will require us to expand our vision beyond horse blinders.

Literature Cited

1. Cline, A.K.; Miler, C.B.; Stewart; G.W.; Wilkinson., J.H, "An Estimate for the Condition Number of a Matrix",, *SIAM J. Numer. Anal.*, V. **16**(2), pp. 368–375, 1977.
2. Kansa, E.J.; Holoborodko," On the ill-conditioned nature of C^{∞} **RBF** strong collocation", E*ng. Anal. Bound. Elem*. vol. **78,** pp. 26–30, 2017.
3. https://www.advanpix.com.
4. Kansa, E.J.; Holoborodko: "Fully and sparsely supported radial basis functions", *Int. J. Comput. Meth. Exper, Measur.,* V. **8**(3)m pp208-219 ,2020.
5. Chen. F.E., *Introduction to Plasma Physics and controlled Fusion,* 3rd edition, ISSN 978-3-319-79391-7, Springer, Heidelberg, 2015.
6. Kirk, D, B, Hwu, W-M.W., *Programming Massively Parallel Computers, Elsevier*, Amsterdam ISBN 978-0- 12-811986-0, 2017.
7. Hilbert, D.: Ein Beitrag zur Theorie des Legendre'schen Polynoms", *Acta Mathematica*, V. **18**: pp.155–159, 1893.
8. Choi, M-D, "Tricks or Treats with the Hilbert Matrix". *The American Mathematical Monthly*., V.**90** (5): pp. 301–312, 1983.
9. Salam.: Conditionality of Linear Systems of Equations and Matrices Using Projective Geometric Algebra, ICCSA 2020 proceedings, Part II, LNCS 12250, pp. 3-17, DOI: 10.1007/978-3-030-58802-1_1, Springer, 2020).
10. Lu, T-T; Shiou, S-H, "Inverses of 2x2 block matrices", Comput. *Math. Applic*., V**43**, pp. 119-129, 2002.
11. Kansa E, Hon Y "Circumventing the ill-conditioning problem with multiquadric radial basis functions: Applications to elliptic partial differential equations" *Computers & Mathematics with Applications*, V.. **39**, (7-8) pp. 123-137, 2000.'
12. kala, V., "Conditionality of linear systems of equations and matrices using projective geometric algebra", ICCSA 2020 Proc., Part II LNCS 12250, pp. 3-17, DOI: 10.1007/987-4-030-53603-2,2, Springer, 2020.