Iain S. Duff

Using FGMRES to obtain backward stability in mixed precision

STFC - Rutherford Appleton Laboratory
Chilton
Didcot
UK OX11 0QX
i.s.duff@rl.ac.uk
Mario Arioli

We are concerned with the solution of

$$Ax \ = \ b, \qquad (1)$$

when A is an $n \times n$ matrix and x and b are vectors of length n. We will solve these systems using a direct method where the matrix A is first factorized as

$$A \to LU,$$

where L and U are triangular matrices. The solution is then obtained through forward elimination

$$Ly = b$$

followed by back substitution

$$Ux = y,$$

where we have omitted permutations required for numerical stability and sparsity preservation for the sake of clarity. When A is symmetric we use an LDL^T factorization where the matrix D is block diagonal with blocks of order 1 and 2, so that we can stably factorize indefinite systems.

On many emerging computer architectures, the use of single precision arithmetic (by which we mean working with 32-bit floating-point numbers) is faster than using double precision. In fact on the Cell processor, single precision working can be more than ten times as fast as using double precision [2]. In addition, half the storage is required when using single precision and the movement of data between memory hierarchies and cache and processing units is much reduced. However, in many applications, a higher accuracy is required than single precision (with a value of machine precision around 10^{-7}) or the matrix can be so ill-conditioned that single precision working is unable to obtain accuracy to even one significant figure .... that is the results are meaningless.

We show how the selective use of double precision post-processing can enable solutions with a backward error (scaled residual) of double precision accuracy (around 10^{-16}) even when the factorization is computed in single precision We show that the use of iterative refinement in double precision may fail when the matrix is ill-conditioned and then show that, even for such badly behaved matrices, the use of FGMRES [3] can produce answers to the desired level of accuracy, namely that the solution process using FGMRES is backward stable at the level of double precision. In [1], we prove that, under realistic assumptions on the matrix and the factorization, the computation in mixed precision, where the LU factorization is computed in single precision and the FGMRES iteration in double precision, gives a backward stable algorithm.

We perform an extensive series of tests using MATLAB on random dense matrices constructed to have specific condition numbers and singular value distribution. We compute the key quantities that appear in our theoretical analysis and show that these support our theory. As one main future development of this work is to study the effective solution of large sparse systems on multicore architectures, we then perform numerical experiments on a set of sparse matrices using a combination of Fortran and MATLAB.

We show the results of runs on some rather ill-conditioned sparse matrices in Table 1. In this case we use restarted FGMRES since the cost of keeping too many vectors can be high for these larger dimensioned systems. We note that although iterative refinement essentially converges on the first three examples, it is still not as accurate as FGMRES and requires many more iterations than FGMRES. On the last example, the convergence of iterative refinement was so slow that we stopped after 53 iterations.

Matrix Id	n	Iterative refinement		FGMRES		
		Total It	$\frac{\|\|b - A\bar{x}_{\hat{k}}\|\|}{(\|\|A\|\| \, \|\|\bar{x}_{\hat{k}}\|\| + \|\|b\|\|)}$	Total It	Inner it	$\frac{\|\|b - A\bar{x}_{\hat{k}}\|\|}{(\|\|A\|\| \, \|\|\bar{x}_{\hat{k}}\|\| + \|\|b\|\|)}$
bcsstk20	485	30	2.1e-15	2	2	1.4e-11
				4	2	3.4e-14
				6	2	7.2e-17
bcsstm27	1224	22	1.6e-15	2	2	5.8e-11
				4	2	1.8e-11
				6	2	6.0e-13
				8	2	1.5e-13
				10	2	1.2e-14
				12	2	2.6e-15
				14	2	1.8e-16
s3rmq4m1	5489	16	2.2e-15	2	2	3.5e-11
				4	2	2.1e-13
				6	2	4.5e-15
				8	2	1.1e-16
s3dkq4m2	90449	53	1.1e-10	10	10	6.3e-17

Table 1: Sparse matrices results.

Bibliography

[1] M. Arioli and I. S. Duff, *Using FGMRES to obtain backward stability in mixed precision*, submitted to ETNA, (2008).

[2] A. Buttari, J. Dongarra, J. Langou, J. Langou, P. Luszczek, and J. Kurzak, *Mixed precision iterative refinement techniques for the solution of dense linear systems*, Int. J. of High Performance Computing Applications **21**(4) (2007), 457–466.

[3] Y.Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Scientific and Statistical Computing **14** (1993), 461–469.