1
Introduction

> To obtain improved convergence rates for the methods of successive displacement we require the coefficient matrix to have a P-condition number as small as possible. If this criterion is not satisfied, then it is advisable to prepare the system or 'precondition' it beforehand.
>
> D. J. Evans. *Journal of the Institute of Mathematics and Applications,* (1968)

> In devising a preconditioner, we are faced with a choice between finding a matrix M that approximates A, and for which solving a system is easier than solving one with A, or finding a matrix M that approximates A^{-1}, so that only multiplication by M is needed.
>
> R. Barrett, *et al.* The *Templates* book. SIAM Publications (1993)

This book is concerned with designing an effective matrix, the so-called preconditioner, in order to obtain a numerical solution *with more accuracy or in less time*. Denote a large-scale linear system of n equations, with $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, by

$$Ax = b \tag{1.1}$$

and one simple preconditioned system takes the following form

$$MAx = Mb. \tag{1.2}$$

(Our primary concern is the real case; the complex case is addressed later.) We shall present various techniques of constructing such a preconditioner M that the preconditioned matrix $A_1 = MA$ has better matrix properties than A. As we see, preconditioning can strike the balance of success and failure of a numerical method.

This chapter will review these introductory topics (if the material proves difficult, consult some suitable textbooks e.g. [80,444] or read the Appendix of [299]).

Section 1.1 Direct and iterative solvers, types of preconditioning
Section 1.2 Norms and condition number
Section 1.3 Perturbation theorems for linear systems and eigenvalues
Section 1.4 The Arnoldi iterations and decomposition
Section 1.5 Clustering, field of values and ϵ-pseudospectrum
Section 1.6 Fast Fourier transforms and fast wavelet transforms
Section 1.7 Numerical solution techniques
Section 1.8 Common theories on preconditioned systems
Section 1.9 Guide to software development and the supplied M-files

Here Sections 1.2–1.5 review some basic theories, Sections 1.6 and 1.7 some numerical tools with which many matrix problems are derived from applied mathematics applications, and Section 1.8 on general discussion of preconditioning issues. Finally Section 1.9 introduces several software Mfiles which the reader can download and use to further the understanding of the underlying topics.

1.1 Direct and iterative solvers, types of preconditioning

There are two types of practical methods for solving the equation (1.1): the direct methods (Chapter 2) and the iterative methods (Chapter 3). Each method produces a numerical solution x, that should approximate the analytical solution $x^* = A^{-1}b$ with a certain number of accurate digits. Modern developments into new solution techniques make the distinction of the two types a bit blurred, because often the two are very much mixed in formulation.

Traditionally, a direct method refers to any method that seeks a solution to (1.1) by simplifying A explicitly

$$Ax = b \qquad \Longrightarrow \qquad A_j x = b_j \qquad \Longrightarrow \qquad Tx = c, \tag{1.3}$$

where T is a much simplified matrix (e.g. T is ideally diagonal) and $c \in \mathbb{R}^n$. The philosophy is essentially *decoupling* the interactions of components of $x = [x_1, \ldots, x_n]^T$ in a new system. One may say that the 'enemy' is A. A somewhat different approach is taken in the Gauss–Purcell method Section 15.5 that views x from a higher space $\mathbb{R}^{n+1}$.

On the other hand, without modifying entries of A, an iterative method finds a sequence of solutions $x_0, x_1, \ldots, x_k, \ldots$ by working closely with the residual vector

$$r = r_k = b - Ax_k, \tag{1.4}$$

which can not only indicate how good x_k is, but also may extract analytical information of matrix A. One hopes that an early termination of iterations will provide a sufficient accurate solution, which is cheaper to obtain than a direct solver. Indeed it is the analytical property of A that determines whether an iterative method will converge at all. See Chapter 3.

A preconditioner may enter the picture of both types of methods. It can help a direct method (Chapter 2) to achieve the maximum number of digits allowable in full machine precision, whenever conditioning is an issue. For an iterative method (Chapter 3), when (often) convergence is a concern, preconditioning is expected to accelerate the convergence to an approximate solution quickly [48,464].

Although we have only listed one typical type of preconditioning in (1.2) namely the *left inverse preconditioner* as the equation makes sense only when $M \approx A^{-1}$ in some sense. There are several other types, each depending on how one intends to approximate A and whether one intends to transform A to $\widetilde{A}$ in a different space before approximation.

Essentially all preconditioners fall into two categories.

Forward Type:	aiming $M \approx A$	
	I (left)	$M^{-1}Ax = M^{-1}b$
	II (right)	$AM^{-1}y = b,\ x = M^{-1}y$
	III (mixed)	$M_2^{-1}AM_1^{-1}y = M_2^{-1}b,\ x = M_1^{-1}y$
Inverse Type:	aiming $M \approx A^{-1}$	
	I (left)	$MAx = Mb$
	II (right)	$AMy = b,\ x = My$
	III (mixed)	$M_2AM_1y = M_2b,\ x = M_1y.$

Clearly matrix splitting type (e.g. incomplete LU decomposition [28] as in Chapter 4) preconditioners fall into the **Forward Type** as represented

$$M^{-1}Ax = M^{-1}b \tag{1.5}$$

while the approximate inverse type preconditioners (e.g. AINV [57] and SPAI [253] as in Chapter 5) fall into the **Inverse Type** and can be represented by (1.2) i.e.

$$MAx = Mb. \tag{1.6}$$

Similarly it is not difficult to identify the same types of preconditioners when A is first transformed into $\widetilde{A}$ in another space, as in Chapters 8–10.

Remark 1.1.1. Here we should make some remarks on this classification.

(1) As with all iterative solution methods, explicit inversion in the **Forward Type** is implemented by a direct solution method e.g. implement $z = M^{-1}w$ as solving $Mz = w$ for z (surely one can use another iterative solver for this).
(2) Each preconditioner notation can be interpreted as a product of simple matrices (e.g. factorized form).
(3) In the context of preconditioning, the notation $M \approx A$ or $M^{-1} \approx A$ should be broadly interpreted as approximating either A (or A^{-1}) directly or simply its certain analytical property (e.g. both M and A have the same small eigenvalues).

1.2 Norms and condition number

The magnitude of any scalar number is easily measured by its modulus, which is non-negative, i.e. $|a|$ for $a \in \mathbb{R}$. The same can be done for vectors in $\mathbb{R}^n$ and matrices in $\mathbb{R}^{m\times n}$, through a non-negative measure called the *norm*. The following definition, using three norm axioms, determines if any such non-negative measure is a norm.

Definition 1.2.2. *Let V be either $\mathbb{R}^n$ (for vectors) or $\mathbb{R}^{m\times n}$ (for matrices). A measure $\|u\|$ of $u \in V$, satisfying the following* **Norm axioms**, *is a valid norm:*

- $\|u\| \geq 0$ *for any u and* $\|u\| = 0$ *is and only if* $u = 0$,
- $\|\alpha u\| = |\alpha|\|u\|$ *for any u and any* $\alpha \in \mathbb{R}$,
- $\|u + v\| \leq \|u\| + \|v\|$ *for any* $u, v \in V$.

Remark that the same axioms are also used for function norms.
One can verify that the following are valid vector norms [80], for $x \in \mathbb{R}^n$,

$$\|x\|_p = \begin{cases} \left(\displaystyle\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \displaystyle\max_{1\leq i\leq n} |x_i|, & \text{if } p = \infty. \end{cases} \tag{1.7}$$

and similarly the following are valid matrix norms, for $A \in \mathbb{R}^{m\times n}$,

$$\|A\|_p = \sup_{x\neq 0\in\mathbb{R}^n} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p=1} \|Ax\|_p, \tag{1.8}$$

where 'sup' denotes 'supremum' and note $Ax \in \mathbb{R}^m$.

While the formulae for vector norms are easy, those for matrices are not. We need to take some specific p in order to present computable formulae from (1.8)

$$
\begin{aligned}
p=1: &\qquad \|A\|_1 = \max_{1\le j\le n}\sum_{i=1}^{m}|a_{ij}| = \max_{1\le j\le n}\|a_j\|_1,\\
p=2: &\qquad \|A\|_2 = \rho(A^TA)^{1/2} = \max_{1\le j\le n}\lambda_j(A^TA)^{1/2} = \sigma_{\max}(A)\\
p=\infty: &\qquad \|A\|_\infty = \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}| = \max_{1\le i\le m}\|\widetilde{a}_i\|_1,
\end{aligned}
\tag{1.9}
$$

where $A\in\mathbb{R}^{m\times n}$ is partitioned in columns first $A=[a_1,\ldots,a_n]$ and then in rows $A=[\widetilde{a}_1^T,\ldots,\widetilde{a}_m^T]^T$, λ is the notation for eigenvalues,[1] ρ denotes the spectral radius and $\sigma_{\max}(A)$ denotes the maximal singular value.[2] Notice that the matrix 2-norm (really an eigenvalue norm) does not resemble the vector 2-norm. The proper counterpart is the following matrix Frobenius norm:

$$
\begin{aligned}
\|A\|_F &= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2} = \sqrt{\sum_{j=1}^{n}\|a_j\|_2^2} = \sqrt{\sum_{i=1}^{m}\|\widetilde{a}_i\|_2^2}\\
&= \left[\mathrm{trace}(A^TA)\right]^{1/2} = \left[\mathrm{trace}(AA^T)\right]^{1/2} = \left[\sum_{j=1}^{n}\lambda_j(AA^T)\right]^{1/2}.
\end{aligned}
\tag{1.10}
$$

A closely related norm is the so-called Hilbert–Schmidt ‘weak’ norm

$$
\|A\|_{HS} = \frac{\|A\|_F}{\sqrt{n}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2}.
$$

We now review some properties of norms for vectors and square matrices.

(1) If $M^{-1}AM=B$, then $A,\ B$ are similar so $\lambda_j(B)=\lambda_j(A)$, although this property is not directly useful to preconditioning since the latter is supposed to change λ.

Proof. For any j, from $Ax_j=\lambda(A)x_j$, we have (define $y_j=M^{-1}x_j$)

$$
M^{-1}AMM^{-1}x_j = \lambda_j(A)M^{-1}x_j \implies By_j=\lambda_j y_j.
$$

Clearly $\lambda_j(B)=\lambda_j(A)$ and the corresponding jth eigenvector for B is y_j.

[1] Recall the definition of eigenvalues λ_j and eigenvectors x_j of a square matrix $B\in\mathbb{R}^{n\times n}$:

$$
Bx_j=\lambda_j x_j, \qquad \text{for } j=1,\cdots,n,
$$

with $\rho(B)=\max_j|\lambda_j|$. Also $det(B)=\prod_{j=1}^n\lambda_j$ and $trace(B)=\sum_{j=1}^n B(j,j)=\sum_{j=1}^n\lambda_j$.

[2] Recall that, if $A=U\Sigma V^T$ with $\Sigma=\mathrm{diag}(\sigma_j)$ for orthogonal U,V, then σ_j's are the singular values of A while U,V contain the left and right singular vectors of A respectively. Denote by $\Sigma(A)$ the spectrum of singular values.

(2) One of the most useful properties of the 2-norm and F-norm is the so-called *norm invariance* i.e. an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ (satisfying $Q^TQ = QQ^T = I$ or $Q^T = Q^{-1}$) does not change the 2-norm of a vector or both the 2-norm and the F-norm of a matrix.
The *proof* is quite simple for the vector case: e.g. let $y = Qx$, then $\|y\|_2^2 = y^Ty = (Qx)^TQx = x^TQ^TQx = x^Tx = \|x\|_2^2$.
For matrices, we use (1.8) to prove $\|PAQ\| = \|A\|$ for F and 2-norms.
Proof. (i) 2-norm case: **(Norm invariance)**
Noting that $\lambda(Q^TA^TAQ) = \lambda(A^TA)$ since Q serves as a similarity transform, then

$$\|PAQ\|^2 = \rho((PAQ)^T\ PAQ) = \rho(Q^TA^TAQ) = \rho(A^TA) = \|A\|^2.$$

(ii) F-norm case: Let $AQ = W = [w_1,\ w_2,\ \ldots,\ w_n]$ (in columns). We hope to show $\|PW\|_F = \|W\|_F$ first. This follows from a column partition of matrices and the previous property for vector norm invariance: $\|PAQ\|_F^2 = \|PW\|_F^2 = \sum_{j=1}^n \|Pw_j\|_2^2 = \sum_{j=1}^n \|w_j\|_2^2 = \|W\|_F^2 = \|AQ\|_F^2$. Next it remains to show that $\|AQ\|_F^2 = \|A\|_F^2$ from a row partition $A = \left(a_1^T\ a_2^T \ldots a_n^T\right)^T$ and again the vector norm invariance:

$$\|AQ\|_F^2 = \sum_{i=1}^n \|a_iQ\|_2^2 = \sum_{i=1}^n \|Q^Ta_i^T\|_2^2 = \sum_{i=1}^n \|a_i^T\|_2^2 = \|A\|_F^2.$$

The same result holds for the complex case when Q is unitary.[3]

(3) The spectral radius is the lower bound for all matrix norms: $\rho(B) \le \|B\|$. The proof for p-norms is quite easy, as (define $Bx_k = \lambda_k x_k$)

$$\rho(B) = \max_{1\le j\le n} |\lambda_j(B)| = |\lambda_k(B)| = \frac{|\lambda_k(B)|\cdot\|x_k\|}{\|x_k\|}$$
$$= \frac{\|Bx_k\|}{\|x_k\|} \le \sup_{x\ne 0}\frac{\|Bx\|_p}{\|x\|_p} = \|B\|_p.$$

For the F-norm, one can use the Schur unitary decomposition

$$B = UTU^H, \qquad \text{with triangular} \qquad T = \Lambda + N, \tag{1.11}$$

where U^H is the conjugate transpose of the matrix U and $\Lambda = \operatorname{diag}(T)$ contains the eigenvalues, so[4] $\|B\|_F^2 = \|T\|_F^2 = \|\Lambda\|_F^2 + \|N\|_F^2$. Hence $\rho(B) \le \sqrt{\sum_{j=1}^n |\lambda_j|^2} \le \|T\|_F$.

[3] In contrast to $Q^TQ = I$ for a real and orthogonal matrix Q, a complex matrix Q is called *unitary* if $Q^HQ = I$.

[4] This equation is used later to discuss non-normality. For a normal matrix, $\|N\| = 0$ as $N = 0$.

(4) If A is symmetric, then $A^T A = A^2$ so the 2-norm is simply: $\|A\|_2 = \rho(A^2)^{1/2} = \rho(A)$.
The proof uses the facts: $\lambda_j(A^2) = \lambda_j(A)^2$ and in general $\lambda_j(A^k) = \lambda_j(A)^k$ for any matrix A which is proved by repeatedly using $Ax_j = \lambda x_j$ (if $\lambda_j \neq 0$ and ∀ integer k).

(5) Matrix ∞-norm (1.9) may be written into a product form

$$\|A\|_\infty = \max_{1\le i\le m} \|\widetilde{a}_i\|_1 = \|\,|A|\mathbf{e}\|_\infty\,, \tag{1.12}$$

with $\mathbf{e} = (1, \ldots, 1)^T$, and similarly

$$\|A\|_1 = \max_{1\le j\le n} \|a_j\|_1 = \left\|\,|A^T|\mathbf{e}\right\|_\infty = \left\|\left(\mathbf{e}^T|A|\right)^T\right\|_\infty,$$

i.e.

$$\|\underbrace{A}_{\text{matrix}}\|_\infty = \|\underbrace{|A|\mathbf{e}}_{\text{vector}}\|_\infty, \qquad \text{and} \qquad \|\underbrace{A}_{\text{matrix}}\|_1 = \|\underbrace{|A^T|\mathbf{e}}_{\text{vector}}\|_\infty.$$

(6) Besides the above 'standard matrix norms', one may also view a $n \times n$ matrix $A \in \mathbb{R}^{n\times n}$ as a member of the long vector space $A \in \mathbb{R}^{n^2}$. If so, A can be measured in vector p-norms directly (note: $\|A\|_{2,v} = \|A\|_F$ so this explains how F-norm is invented)

$$\|A\|_{p,v} = \left(\sum_{j=1}^{n}\sum_{k=1}^{n} |a_{jk}|^p\right)^{1/p}.$$

(7) Different norms of the same quantity can only differ by at most a constant multiple that depends on the dimension parameter n. For example,

$$\|A\|_1 \le n\|A\|_\infty, \quad \|A\|_\infty \le n\|A\|_1, \qquad \text{and} \tag{1.13}$$

$$\|A\|_2 \le \sqrt{n}\|A\|_1, \quad \|A\|_1 \le \sqrt{n}\|A\|_2, \tag{1.14}$$

from the vector norm inequalities

$$\frac{\|x\|_1}{\sqrt{n}} \le \|x\|_2 \le \|x\|_1. \tag{1.15}$$

Proof. To prove (1.13), use (1.8) and matrix partitions in (1.9): for any $x \neq 0 \in \mathbb{R}^n$,

$$\frac{\|Ax\|_1}{\|x\|_1} = \frac{\sum_{i=1}^n |\widetilde{a}_i x|}{\sum_{i=1}^n |x_i|} \le \frac{n \max_i |\widetilde{a}_i x|}{\sum_{i=1}^n |x_i|} \le \frac{n \max_i |\widetilde{a}_i x|}{\max_i |x_i|} = \frac{n\|Ax\|_\infty}{\|x\|_\infty},$$

$$\frac{\|Ax\|_\infty}{\|x\|_\infty} = \frac{n \max_i |\widetilde{a}_i x|}{n \max_i |x_i|} \le \frac{n \max_i |\widetilde{a}_i x|}{\sum_{i=1}^n |x_i|} \le \frac{n \sum_{i=1}^n |\widetilde{a}_i x|}{\sum_{i=1}^n |x_i|} = \frac{n\|Ax\|_1}{\|x\|_1}.$$

Thus the proof for (1.13) is complete from (1.8). To prove the left inequality (since the second one is easy) in (1.15), we can use induction and the Cauchy–Schwarz inequality

$$|x \cdot y| \le \|x\|_2 \|y\|_2, \quad ||x_1| + |x_2|| \le \sqrt{2}\sqrt{x_1^2 + x_2^2}.$$

Using (1.15), inequalities (1.14) follow from (1.8). ■

Further results will be given later when needed; consult also [80,229,280].

(8) The condition number of a nonsingular matrix $A \in \mathbb{R}^{n\times n}$ is defined as

$$\kappa(A) = \|A\|\|A^{-1}\|. \tag{1.16}$$

From $\|AB\| \le \|A\|\|B\|$ and $\|I\| = 1$, we have

$$\kappa(A) \ge 1. \tag{1.17}$$

Here certain matrix norm is used and sometimes for clarity we write explicitly $\kappa_\ell(A)$ if the ℓ-norm is used (e.g. $\ell = 1, 2$ or $\ell = F$).

Although widely publicized and yet vaguely convincing, a condition number measures how well-conditioned a matrix is. However, without involving a particular context or an application, the meaning of $\kappa(A)$ can be seen more precisely from the following two equivalent formulae [280]:

$$\kappa(A) = \lim_{\epsilon\to 0} \sup_{\|\Delta A\|\le\epsilon\|A\|} \frac{\|(A+\Delta A)^{-1} - A^{-1}\|}{\epsilon\|A^{-1}\|}, \tag{1.18}$$

$$\kappa(A) = \frac{1}{\operatorname{dist}(A)} \qquad \text{with} \quad \operatorname{dist}(A) = \min_{A+\Delta A \text{ singular}} \frac{\|\Delta A\|}{\|A\|}. \tag{1.19}$$

If A is 'far' from its nearest singular neighbour, the condition number is small.

Two further issues are important. Firstly, if A is nonsingular and P, Q are orthogonal matrices, then $\kappa(PAQ) = \kappa(A)$ for F and 2-norms.
Proof. This result is a consequence of matrix norm invariance:

$$\kappa(PAQ) = \|(PAQ)^{-1}\|\,\|PAQ\| = \|Q^T A^{-1} P^T\|\,\|A\| = \|A^{-1}\|\,\|A\| = \kappa(A).$$

Secondly, if A is symmetric and nonsingular, then eigensystems are special so

$$\kappa_2(A) = \|A\|_2\|A^{-1}\|_2 = \rho(A)\rho(A^{-1}) = \frac{|\lambda|_{\max}}{|\lambda|_{\min}}. \tag{1.20}$$

For SPD matrices, as $\lambda_j(A) > 0$,

$$\kappa_2(A) = \frac{\lambda_{\max}}{\lambda_{\min}}. \tag{1.21}$$

Example 1.2.3. *The following* 3×3 *matrix*

$$A = \begin{pmatrix} 12 & 1 & -17 \\ 5 & 4 & -9 \\ 7 & 1 & -10 \end{pmatrix}$$

has three eigenvalues $\lambda = 1, 2, 3,$ *from solving* $|A - \lambda I| = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$, *corresponding to three eigenvectors* $x_1 = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$, $x_2 = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$ *and* $x_3 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$.

Clearly the spectral radius of A is $\rho(A) = 3$, $\|A\|_1 = 36$ *and* $\|A\|_\infty = 30$. *Finally putting the eigensystems* $Ax_j = \lambda_j x_j$ *into matrix form, noting* $X = [x_1\ x_2\ x_3]$ *and* $AX = [Ax_1\ Ax_2\ Ax_3]$, *we obtain* $AX = XD$ *with* $D = diag(1, 2, 3)$. *From* $A^{-1} = \begin{pmatrix} -31/6 & -7/6 & 59/6 \\ -13/6 & -1/6 & 23/6 \\ -23/6 & -5/6 & 43/6 \end{pmatrix}$, *we can compute that* $\kappa_1(A) = 36\ (125/6) = 750$.

Remark 1.2.4. Although this book will focus on fast solvers for solving the linear system (1.1), the solution of the eigenvalue problems may be sought by techniques involving indefinite and singular linear systems and hence our preconditioning techniques are also applicable to ‘preconditioning eigenvalue problems’ [36,58,360]. Refer also to [486].

1.3 Perturbation theories for linear systems and eigenvalues

A computer (approximate) solution will not satisfy (1.1) exactly so it is of interest to examine the perturbation theory.

Theorem 1.3.5. *Let* $Ax = b$ *and* $(A + \Delta A)(x + \Delta x) = b + \Delta b$.
If $\|A^{-1}\|\|\Delta A\| < 1$, *then*

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\kappa(A)}{1 - \|A^{-1}\|\|\Delta A\|}\left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right). \tag{1.22}$$

The *proof* of this result is easy once one notes that

$$\Delta x = (A + \Delta A)^{-1}(-\Delta A x + \Delta b) = \left(I + A^{-1}\Delta A\right)^{-1} A^{-1}(-\Delta A x + \Delta b),$$

$\| \left(I + A^{-1}\Delta A\right)^{-1} \| \le 1/(1 - \|A^{-1}\| \|\Delta A\|), \ \|b\| \le \|A\| \|x\|.$

Here for our application to iterative solution, ΔA is not a major concern but Δb will be reflected in the usual stopping criterion based on residuals. Clearly the condition number $\kappa(A)$ will be crucial in determining the final solution accuracy as computer arithmetic has a finite machine precision.

Eigenvalue perturbation. We give a brief result on eigenvalues simply for comparison and completeness [218].

Theorem 1.3.6. (Bauer–Fike). *Let $A, X \in \mathbb{R}^{n\times n}$ and X be nonsingular. For $D = diag(\lambda_i)$ with $\lambda_i = \lambda_i(A)$, let $\|.\|$ be one of these norms: 1, 2 or ∞, such that $\|D\| = \max_{1\le i\le n} |\lambda_i|$. If $B = A + \Delta A$ is a perturbed matrix of A and $X^{-1}AX = D$, then all eigenvalues of B are inside the union of n discs $\bigcup_{i=1}^{n} \Omega_i$, where the discs are defined by*

General case: 1, 2, ∞ norm and any matrix X: $cond(X) = \|X\| \, \|X^{-1}\| \ge 1$

$$\Omega_i = \{z \in C \quad | \quad |z - \lambda_i| \le cond(X)\|\Delta A\|\}.$$

Special case: 2-norm and orthogonal matrix X: $cond(X) = \|X\|_2 \, \|X^{-1}\|_2 = 1$

$$\Omega_i = \{z \in C \quad | \quad |z - \lambda_i| \le \|\Delta A\|_2\}.$$

Proof. Define $K = D - \mu I$. If μ is an eigenvalue of B, then $B - \mu I = A + \Delta A - \mu I$ is singular; we may assume that this particular $\mu \ne \lambda_i(A)$ for any i (otherwise the above results are already valid as the difference is zero). That is, we assume K is nonsingular. Also define $W = K^{-1}X^{-1}\Delta A X$. Now consider decomposing

$$\begin{aligned} B - \mu I = A + \Delta A - \lambda I &= X D X^{-1} + \Delta A - \mu I \\ &= X\left[D + X^{-1}\Delta A X - \mu I\right] X^{-1} \\ &= X\left[(D - \mu I) + X^{-1}\Delta A X\right] X^{-1} \\ &= X\left[K + X^{-1}\Delta A X\right] X^{-1} \\ &= X K \left[I + K^{-1}X^{-1}\Delta A X\right] X^{-1} \\ &= X K \left[I + W\right] X^{-1}. \end{aligned}$$

Clearly, by taking determinants both sides, matrix $I + W$ must be singular as $B - \mu I$ is, so W has the eigenvalue -1. Further from the property