Perturbation Analysis of Higham Squared Maximum Growth Matrices

a dissertation presented
by
Bowen Zhu
to
The Department of Applied Computation

in partial fulfillment of the requirements
for the degree of
Master of Engineering
in the subject of
Computational Science and Engineering

Harvard University
Cambridge, Massachusetts
May 2024

Thesis advisor: Professor Alan Edelman Bowen Zhu

Perturbation Analysis of Higham Squared Maximum Growth Matrices

Abstract

This thesis examines the behavior of Higham^2 matrices in Gaussian elimination through perturbation analysis. Higham^2 matrices, including the special case of Wilkinson matrices, are known for achieving the maximum growth under partial pivoting. Through theoretical analysis and numerical experiments, this thesis highlights the sensitivity of these matrices to perturbations and how these small perturbations can be strategically applied to matrix entries to reduce the growth, thus enhancing computational stability and accuracy.[1]

[1] All code is made available at `https://github.com/Bowen1Zhu/growth_factor`.

Contents

Listing of figures

Acknowledgments

The author would like to thank Dr. Alan Edelman, Dr. John Urschel, and Dr. Daniel Weinstock for their gracious guidance, help, and support throughout the development and writing of this thesis.

0
Introduction

The solution of a linear system $Ax = b$ is one of the oldest problems in mathematics. One of the most fundamental and important techniques for solving a linear system is Gaussian elimination, in which a matrix is factored into the product of a lower and upper triangular matrix. Given an $n \times n$ matrix A, Gaussian elimination performs a sequence of rank-one transformations, resulting in the sequence of matrices $A^{(k)} \in \mathbb{C}^{k\times k}$ for k equals n to 1, satisfying

$$A^{(k)} = M^{(2,2)} - M^{(2,1)}[M^{(1,1)}]^{-1}M^{(1,2)}, \qquad \text{where } A = \begin{array}{cc} \begin{array}{cc} n-k & k \end{array} & \\ \begin{bmatrix} M^{(1,1)} & M^{(1,2)} \\ M^{(2,1)} & M^{(2,2)} \end{bmatrix} & \begin{array}{c} n-k \\ k \end{array} \end{array}.$$

The resulting LU factorization of A is encoded by the first row and column of each of the iterates $A^{(k)}$, $k = 1, \ldots, n$. Not all matrices have an LU factorization, and a permutation of the rows (or columns) of the matrix may be required. In addition, performing computations in finite precision can elicit issues due to round-off error. The error due to rounding in Gaussian elimination for a matrix A in some fixed precision is controlled by the growth factor of the Gaussian elimination algorithm, defined by

$$g(A) := \frac{\max_k |A^{(k)}|_\infty}{|A|_\infty},$$

where $|\cdot|_\infty$ is the entry-wise matrix infinity norm (see[5] Theorem 3.3.1 for details). For this reason, understanding the growth factor is of both theoretical and practical importance.

Partial pivoting is the most popular method for performing Gaussian elimination, and produces a factorization $PA = LU$, where, at each step of Gaussian elimination applied to PA, the pivot is the largest magnitude entry of the first column. This method is widely available through interfaces from high level languages such as Julia[1], Mathematica[19], Matlab[15], Python NumPy[7], R[11] etc. Many researchers have studied and continue to study the question of why Gaussian elimination with partial pivoting has been so very effective[4,8,9,10,13,14,16,17]. In contrast to complete pivoting, where the existence of matrices with even super-linear growth remains an open problem[2,3,6], it has been known since Wilkinson's classic text *The Algebraic Eigenvalue Problem*[18] p.212 that, for partial pivoting, the growth factor is bounded above by 2^{n-1} and that this quantity can be achieved by the matrix

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 & 1 \\ -1 & \cdots & -1 & 1 & 1 \\ -1 & \cdots & -1 & -1 & 1 \end{pmatrix}. \tag{0.0.1}$$

Much later, Higham and Higham (from now on denoted Higham[2]) identified the complete set of n by n real matrices that achieve the maximal growth of 2^{n-1}[9]. We call such matrices Higham^2 matrices (see Proposition 0.0.1 for a description).

Proposition 0.0.1. [9] *Theorem 2.2 Every matrix $A \in \mathrm{GL}_n(\mathbb{R})$, $\|A\|_{max} = 1$, with growth factor under partial pivoting equal to 2^{n-1} must be of the form*

$$DPA = \begin{pmatrix} L & 0 \\ -\mathbf{1}^T & 1 \end{pmatrix} \begin{pmatrix} U & \boldsymbol{u} \\ 0 & 2^{n-1} \end{pmatrix} = \begin{pmatrix} LU & \mathbf{1} \\ -\mathbf{1}^T U & 1 \end{pmatrix}, \tag{0.0.2}$$

where $D, P \in \mathrm{GL}_n(\mathbb{R})$ is a ± 1 diagonal matrix and permutation matrix, respectively, $L \in \mathrm{GL}_{n-1}(\mathbb{R})$ is lower uni-triangular with $L_{ij} = -1$ for all $i > j$, $\boldsymbol{u} = (1, 2, \ldots, 2^{n-2})^T \in \mathbb{R}^{n-1}$, and $U \in \mathrm{GL}_{n-1}(\mathbb{R})$ is upper triangular, with entries satisfying $\|LU\|_{max} \leq 1$ and $\|\mathbf{1}^T U\|_\infty \leq 1$.

Partial pivoting is practically unstable in the worst case. For partial pivoting, open questions largely concern the behavior of worst-case instances under random perturbation (i.e., smoothed analysis, see [12,13]) and of random matrices (see [17,10]). Here, we study how stable large growth in partial pivoting is under small perturbations. We consider the class of partially pivoted matrices $A \in \mathbb{R}^{n \times n}$ with maximal growth $g(A) = 2^{n-1}$, characterized by Higham and Higham[9].

A scalar quantity of interest is the last pivot of a Higham^2 matrix, which is a differentiable func-

[2] We denote "Higham and Higham" as Higham^2, to be read as "Higham squared," yet we realize this has the appearance of a footnote, so for readers who saw it this way, we have included this footnote.

tion of the matrix entries as Higham^2 matrices can not be singular. We can therefore ask for the gradient of this last pivot or, even better, to have a full (non-infinitesimal) perturbation analysis of the last pivot (for Gaussian elimination without pivoting). We provide such a perturbation analysis in Theorem 1.2.1. The last pivot is an ideal quantity to measure in order to understand the growth factor, as every entry of U is the last pivot of the LU factorization of some submatrix of A. We observe that generically, large growth does not last very long in the sense that often a small perturbation can dramatically reduce a large pivot. We have a mental image that the Higham^2 matrices live on a kind of "ridge" that one can easily fall off of. This picture is consistent with the smoothed analysis of Sankar, Spielman, and Teng[13]. The structure of the Higham^2 matrices provides an ideal setting to better understand the ridge and its profile. Perhaps unsurprisingly, not all directions of descent are created equal. We provide numerical experiments (in Chapter 2) to visualize the effects of perturbing Higham^2 matrices and confirm the conclusions gleaned from the theoretical results of Theorem 1.2.1.

1

Entrywise Perturbation Analysis

1.1 General Matrices

Here we provide mathematical estimates for the effects of entrywise perturbations on the last pivot of the LU factorization of a matrix (Lemma 1.1.1), recall an explicit representation of Higham[2] matrices (Proposition 0.0.1), and consider the effects of entrywise perturbations for this class (Theorem 1.2.1). The theoretical results of Theorem 1.2.1 give insight into the experimental results we observe in Chapter 2.

Lemma 1.1.1. *Let*

$$A = \begin{pmatrix} L & 0 \\ \boldsymbol{\ell}^T & 1 \end{pmatrix} \begin{pmatrix} U & \boldsymbol{u} \\ 0 & p \end{pmatrix} \in \mathrm{GL}_n(\mathbb{R}),$$

where $L \in \mathrm{SL}_{n-1}(\mathbb{R})$ *is lower unitriangular,* $U \in \mathrm{GL}_{n-1}(\mathbb{R})$ *is upper triangular, and* $\boldsymbol{\ell}, \boldsymbol{u} \in \mathbb{R}^{n-1}$*. Then the LU factorization, if it exists, of* $A + \varepsilon\, \boldsymbol{e}_i \boldsymbol{e}_j^T$*, where* $\boldsymbol{e}_i$ *is the* i^{th} *standard basis vector, has last pivot* $p_\varepsilon^{(i,j)}$*, where*

$$p_\varepsilon^{(i,j)} = \begin{cases} p + \varepsilon \dfrac{(U^{-1}\boldsymbol{u})_j (\boldsymbol{\ell}^T L^{-1})_i}{1 + \varepsilon (LU)^{-1}_{ji}} & \text{for} \quad i, j < n \\ p - \varepsilon (\boldsymbol{\ell}^T L^{-1})_i & \text{for} \quad i < n,\ j = n \\ p - \varepsilon (U^{-1}\boldsymbol{u})_j & \text{for} \quad i = n,\ j < n \\ p + \varepsilon & \text{for} \quad i = j = n \end{cases} . \tag{1.1.1}$$

Proof. The matrices A and A^{-1} have block form

$$A = \begin{pmatrix} L & 0 \\ \boldsymbol{\ell}^T & 1 \end{pmatrix} \begin{pmatrix} U & \boldsymbol{u} \\ 0 & p \end{pmatrix} = \begin{pmatrix} LU & L\boldsymbol{u} \\ \boldsymbol{\ell}^T U & p + \boldsymbol{\ell}^T \boldsymbol{u} \end{pmatrix}$$

and

$$A^{-1} = \begin{pmatrix} U^{-1} & -p^{-1}U^{-1}\boldsymbol{u} \\ 0 & p^{-1} \end{pmatrix} \begin{pmatrix} L^{-1} & 0 \\ -\boldsymbol{\ell}^T L^{-1} & 1 \end{pmatrix} = \begin{pmatrix} (LU)^{-1} + p^{-1}U^{-1}\boldsymbol{u}\boldsymbol{\ell}^T L^{-1} & -p^{-1}U^{-1}\boldsymbol{u} \\ -p^{-1}\boldsymbol{\ell}^T L^{-1} & p^{-1} \end{pmatrix}.$$

Let us first consider the case $i, j < n$. We have

$$p_\varepsilon^{(i,j)} = \frac{\det(A + \varepsilon\, \boldsymbol{e}_i \boldsymbol{e}_j^T)}{\det(LU + \varepsilon\, \hat{\boldsymbol{e}}_i \hat{\boldsymbol{e}}_j^T)} = \frac{\det(A)}{\det(LU)} \left(\frac{1 + \varepsilon \boldsymbol{e}_j^T A^{-1} \boldsymbol{e}_i}{1 + \varepsilon \hat{\boldsymbol{e}}_j^T (LU)^{-1} \hat{\boldsymbol{e}}_i} \right) = p \left(\frac{1 + \varepsilon A_{ji}^{-1}}{1 + \varepsilon (LU)_{ji}^{-1}} \right),$$

where $\hat{\boldsymbol{e}}_i$ is the i^{th} standard basis vector in $\mathbb{R}^{n-1}$. For $i, j < n$,

$$A_{ji}^{-1} = (LU)_{ji}^{-1} + p^{-1} (U^{-1} \boldsymbol{u})_j (\boldsymbol{\ell}^T L^{-1})_i,$$

and so

$$p_\varepsilon^{(i,j)} - p = p \left[\frac{1 + \varepsilon A_{ji}^{-1}}{1 + \varepsilon (LU)_{ji}^{-1}} - 1 \right] = \frac{\varepsilon p (A_{ji}^{-1} - (LU)_{ji}^{-1})}{1 + \varepsilon (LU)_{ji}^{-1}} = \frac{\varepsilon (U^{-1} \boldsymbol{u})_j (\boldsymbol{\ell}^T L^{-1})_i}{1 + \varepsilon (LU)_{ji}^{-1}}.$$

When $i = n$ or $j = n$,

$$p_\varepsilon^{(i,j)} = \frac{\det(A + \varepsilon\, \boldsymbol{e}_i \boldsymbol{e}_j^T)}{\det(LU)} = p(1 + \varepsilon A_{ji}^{-1}).$$

Noting that $A_{jn}^{-1} = -p^{-1} (U^{-1} \boldsymbol{u})_j$ for $j < n$, $A_{ni}^{-1} = -p^{-1} (\boldsymbol{\ell}^T L^{-1})_i$ for $i < n$, and $A_{nn}^{-1} = p^{-1}$ completes the proof. $\square$

1.2 Higham² Matrices

Applying Lemma 1.1.1 to Higham² matrices (described in Proposition 0.0.1) gives the following theorem.

Theorem 1.2.1. *Let A, $\|A\|_{max} = 1$, be a Higham² matrix of the form in Equation 0.0.2 (e.g., as described in Proposition 0.0.1 with $P = D = I$). Then the LU factorization, if it exists, of $A + \varepsilon\, \boldsymbol{e}_i \boldsymbol{e}_j^T$,* *where $\boldsymbol{e}_i$ is the i^{th} standard basis vector, has last pivot*

$$p_\varepsilon^{(i,j)} = \begin{cases} \dfrac{1}{2^{1-n} + \varepsilon\, 2^{-i} \sum_{\ell=1}^{n-j} 2^{-\ell} U^{-1}_{j,n-\ell}} & \textit{for} \quad i < j < n \\ \dfrac{1 + \frac{1}{2}\varepsilon \left(U^{-1}_{ji} - \sum_{\ell=1}^{i-j} 2^{-\ell} U^{-1}_{j,i-\ell}\right)}{2^{1-n} + \varepsilon(2^{1-n} U^{-1}_{ji} + 2^{-i} \sum_{\ell=1}^{n-i-1} 2^{-\ell} U^{-1}_{j,n-\ell})} & \textit{for} \quad j \le i < n \\ 2^{n-1}\left(1 + \varepsilon 2^{-i}\right) & \textit{for} \quad i < j = n \\ 2^{n-1}\left(1 + \varepsilon \displaystyle\sum_{\ell=1}^{n-j} 2^{-\ell} U^{-1}_{j,n-\ell}\right) & \textit{for} \quad j < i = n \\ 2^{n-1} + \varepsilon & \textit{for} \quad i = j = n \end{cases}, \tag{1.2.1}$$

Proof. By Proposition 0.0.1, $A = \begin{pmatrix} L & 0 \\ -\mathbf{1}^T & 1 \end{pmatrix} \begin{pmatrix} U & \boldsymbol{u} \\ 0 & 2^{n-1} \end{pmatrix}$, where $L \in \mathrm{SL}_{n-1}(\mathbb{R})$ is lower uni-triangular with $L_{ij} = -1$ for all $i > j$, $\boldsymbol{u} = (1, 2, ..., 2^{n-2})^T$, and $U \in \mathrm{GL}_{n-1}(\mathbb{R})$ is upper triangular. The matrix L has a simple structure, and its inverse has entries given by $L^{-1}_{ij} = \varphi(i-j)$, where $\varphi(k)$ equals zero for $k < 0$, one for $k = 0$, and 2^{k-1} for $k > 0$. Therefore,

$$(\mathbf{1}^T L^{-1})_i = \sum_{k=i}^{n-1} \varphi(k-i) = 2^{n-i-1} \qquad \text{and} \qquad (LU)^{-1}_{ji} = \sum_{k=i}^{n-1} \varphi(k-i) U^{-1}_{jk} = U^{-1}_{ji} + \sum_{k=i+1}^{n-1} 2^{k-i-1} U^{-1}_{jk}.$$

Applying Lemma 1.1.1 to A and noting that $(U^{-1}\boldsymbol{u})_j = \sum_{k=j}^{n-1} 2^{k-1} U^{-1}_{jk}$, we have $p_\varepsilon^{(n,n)} = 2^{n-1} + \varepsilon$, $p_\varepsilon^{(i,n)} = 2^{n-1}(1 + \varepsilon\, 2^{-i})$ for $i < n$, $p_\varepsilon^{(n,j)} = 2^{n-1}(1 + \varepsilon \sum_{\ell=1}^{n-j} 2^{-\ell} U^{-1}_{j,n-\ell})$ for $j < n$, and

$$p_\varepsilon^{(i,j)} = 2^{n-1} - \frac{\varepsilon 2^{n-i-1} \sum_{k=j}^{n-1} 2^{k-1} U^{-1}_{jk}}{1 + \varepsilon(U^{-1}_{ji} + \sum_{k=i+1}^{n-1} 2^{k-i-1} U^{-1}_{jk})} \qquad \text{for} \quad i, j < n.$$

When $i < j < n$, $U^{-1}_{ji} = 0$ and $\sum_{k=i+1}^{n-1} 2^{k-i-1} U^{-1}_{jk} = \sum_{k=j}^{n-1} 2^{k-i-1} U^{-1}_{jk}$, and so $p_\varepsilon^{(i,j)}$ equals

$(2^{1-n} + \varepsilon\, 2^{-i} \sum_{\ell=1}^{n-j} 2^{-\ell} U^{-1}_{j,n-\ell})^{-1}$. Finally, in the case $j \leq i < n$, we have

$$p_\varepsilon^{(i,j)} = \frac{1 + \frac{1}{2}\varepsilon\,(U^{-1}_{ji} - \sum_{\ell=1}^{i-j} 2^{-\ell} U^{-1}_{j,i-\ell})}{2^{1-n} + \varepsilon(2^{1-n} U^{-1}_{ji} + 2^{-i} \sum_{\ell=1}^{n-i-1} 2^{-\ell} U^{-1}_{j,n-\ell})}.$$

$\square$

1.3 Implications

Equation 1.2.1 of Theorem 1.2.1 deserves a number of observations. Let us restrict our attention to the case $i < j$. We note that, as long as the quantity $\varepsilon\, 2^{-i} \sum_{\ell=1}^{n-j} 2^{-\ell} U^{-1}_{k,n-\ell}$ is not exponentially small in n, the last pivot is roughly equal to $2^i\, (\varepsilon \sum_{\ell=1}^{n-j} 2^{-\ell} U^{-1}_{k,n-\ell})^{-1}$. In general, one would expect this to be the case for "most" Higham^2 matrices when ε is only polynomially small, although exceptions certainly exist (e.g., the Wilkinson matrix $U = I$ from Equation 0.0.1, and others). This intuition is supported by experimental results in Chatper 2. The case $i = 1$ and $j = n-1$ is particularly striking, as

$$|U_{n-1,n-1}| = \left| \frac{(LU)_{n-1,n-1} + (\mathbf{1}^T U)_{n-1}}{2} \right| \leq 1,$$

giving

$$p_\varepsilon^{(1,n-1)} = \frac{4 U_{n-1,n-1}}{\varepsilon + 2^{-(n-3)} U_{n-1,n-1}} = \frac{U_{n-1,n-1}}{\varepsilon} \left(4 - \frac{2^{-(n-5)} U_{n-1,n-1}}{\varepsilon + 2^{-(n-3)} U_{n-1,n-1}} \right),$$

and so

$$|p_\varepsilon^{(1,n-1)}| \leq \frac{4 + o_n(1)}{|\varepsilon|}, \qquad \text{where} \qquad |o_n(1)| < 2^{-(n-6)} \text{ for } |\varepsilon| > 2^{-(n-4)}.$$

The entry $1, n-1$ is an example of a perturbation direction that will always produce a small last pivot when ε is only polynomially small. Finally, we note that, while Lemma 1.1.1 and Theorem

1.2.1 apply only to the last pivot, this framework holds for arbitrary entries of U, as every entry of U is the last pivot of the LU factorization of some submatrix of A. In Section 5.1, we make use of the insights gained from Theorem 1.2.1 to suggest perturbations tailored to the most influential components of A.

2

Numerical Observations of Entrywise Perturbations

In this chapter, we explore empirical findings on how perturbations at various entries affect the last pivot value in both general Higham^2 matrices and the Wilkinson matrix. We first present Figure 2.1 which illustrates the overall impact of these perturbations on uniformly random Higham^2 matrices and the Wilkinson matrix, providing a visual summary of our key results. We then analyze these effects based on our empirical observations.

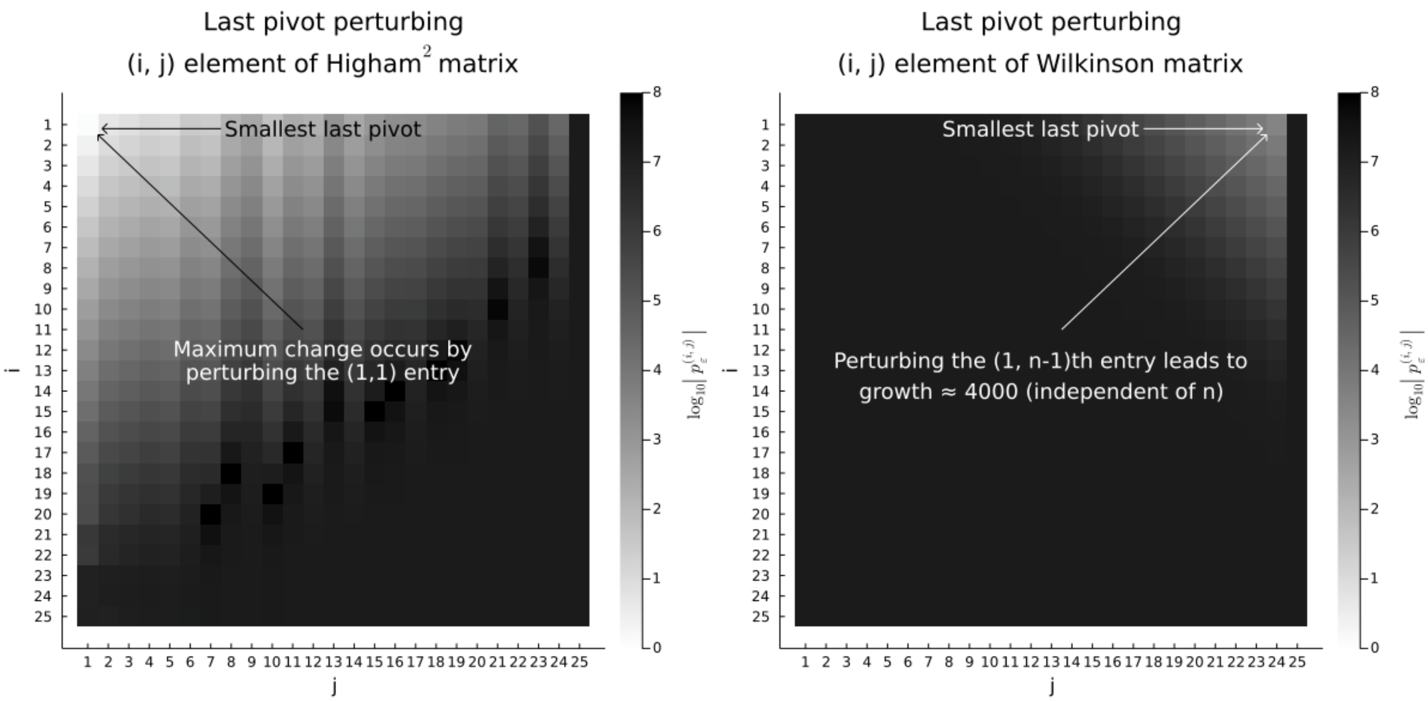

Figure 2.1: Heatmaps of the effect of perturbations at different entries in a random 25×25 Higham^2 and Wilkinson matrix.

The upper plot presents a heatmap of the log ratios $\log \left|p_\varepsilon^{(i,j)}\right|$ for a random 25×25 Higham^2 matrix, normalized such that $\|A\|_{max} = 1$. The lower plot presents a similar heatmap of the log ratios $\log \left|p_\varepsilon^{(i,j)}\right|$ of the 25×25 Wilkinson matrix. Each (i, j)-th grid corresponds to the log ratio $\log \left|p_\varepsilon^{(i,j)}\right|$ at $\varepsilon = 1e - 3$.

In the Higham^2 matrix, the most significant reduction in the last pivot occurs when perturbing the $(1, 1)$ entry, achieving a value of 1.68×10^7. Perturbations in the upper left entries typically lead to substantial reductions.

Conversely, in the Wilkinson matrix, the reduction effects are comparatively smaller, with the most significant reduction occurring at the $(1, n - 1)$ entry, resulting in a pivot value of approximately 4000 (regardless of the matrix size n).

Perturbations in the upper right entries (excluding the last column) tend to yield relatively significant reductions.

Figure 2.1 demonstrates that in uniformly sampled Higham^2 matrices, perturbing the upper left entries typically results in the maximum reduction of the last pivot values. However, in the special case of the Wilkinson matrix, this overall reduction effect is relatively less significant, and perturbing the upper right entries (excluding the rightmost column) reduces the last pivot values most effectively. In the following sections, we will provide an analysis of these phenomena.

2.1 Perturbation of the Largest Reduction

Since U is upper triangular, its inverse U^{-1} is also upper triangular. Thus, the following lemma ensues.

Lemma 2.1.1. *For any lower triangular entry (a, b) with $a > b$, the ratio* $\dfrac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{ab}}{(U^{-1}L^{-1})_{ab}} = -p.$

Proof. Since

$$(\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{jk} = -2^{j-1-k}p,$$

and

$$L^{-1}_{jk} = \begin{cases} 0 & j < k \\ 1 & j = k \\ 2^{j-1-k} & j > k \end{cases},$$

then

$$\begin{aligned}
(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{ab} &= \sum_{c=1}^{n} U^{-1}_{ac}(\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{cb} \\
&= \sum_{c=1}^{a-1} 0 \cdot (\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{cb} + \sum_{c=a}^{n} U^{-1}_{ac}(\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{cb} \\
&= \sum_{c=a}^{n} U^{-1}_{ac} \cdot (-2^{c-1-b}p)
\end{aligned}$$

and

$$
\begin{aligned}
(U^{-1}L^{-1})_{ab} &= \sum_{c=1}^{n} U^{-1}_{ac} L^{-1}_{cb} \\
&= \sum_{c=1}^{a-1} 0 \cdot L^{-1}_{cb} + \sum_{c=a}^{n} U^{-1}_{ac} L^{-1}_{cb} \\
&= \sum_{c=a}^{n} U^{-1}_{ac} \cdot 2^{c-1-b}.
\end{aligned}
$$

Thus we have

$$
\frac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{ab}}{(U^{-1}L^{-1})_{ab}} = \frac{\sum_{c=a}^{n} U^{-1}_{ac} \cdot (-2^{c-1-b}p)}{\sum_{c=a}^{n} U^{-1}_{ac} \cdot 2^{c-1-b}} = -p \quad \text{for } a > b.
$$

□

Claim 2.1.2. *For a uniformly sampled Higham*[2] *matrix with a sufficiently large size, $p_\varepsilon^{(1,1)}$ tends to be small.*

The above statement is reflected in Figure 2.1. This assertion can be derived from the fact that U^{-1}_{11} tends to be the smallest element in magnitude in the first row of U^{-1}, so in general

$$
\frac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{11}}{(U^{-1}L^{-1})_{11}} = \frac{U^{-1}_{11} \cdot (-2^{1-2}p) + \sum_{c=2}^{n} U^{-1}_{1,c} \cdot (-2^{c-2}p)}{U^{-1}_{11} + \sum_{c=2}^{n} U^{-1}_{1,c} \cdot 2^{c-2}} \approx -p,
$$

and given n sufficiently large,

$$
p_\varepsilon^{(1,1)} = p + \varepsilon \frac{(U^{-1}\boldsymbol{u})_1(\boldsymbol{\ell}^T L^{-1})_1}{1 + \varepsilon(LU)^{-1}_{11}} \approx p + \varepsilon \frac{(U^{-1}\boldsymbol{u})_1(\boldsymbol{\ell}^T L^{-1})_1}{\varepsilon(LU)^{-1}_{11}} \approx p + \varepsilon \frac{-p}{\varepsilon} \approx 0.
$$

However, there is no guarantee that $p_\varepsilon^{(1,1)}$ is always small, as we can manually construct cases where $\left|U^{-1}_{11}\right|$ is not relatively small (e.g., the Wilkinson matrix) or $\left|\varepsilon(LU)^{-1}_{11}\right| \not\gg 1$. In fact,

$$
\begin{aligned}
p_\varepsilon^{(1,1)} &= p\left(1+\frac{\varepsilon}{p}\,\frac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{11}}{1+\varepsilon(U^{-1}L^{-1})_{11}}\right)\\
&= p\left(1+\frac{\varepsilon}{p}\,\frac{U_{11}^{-1}\cdot(-2^{1-2}p)+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot(-2^{c-2}p)}{1+\varepsilon(U_{11}^{-1}+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot 2^{c-2})}\right)\\
&= p\left(1-\frac{U_{11}^{-1}\cdot 2^{1-2}+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot 2^{c-2}}{\frac{1}{\varepsilon}+U_{11}^{-1}+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot 2^{c-2}}\right)\\
&= (2^{n-1}\theta)\cdot\left(\frac{\frac{1}{\varepsilon}+\frac{U_{11}^{-1}}{2}}{\frac{1}{\varepsilon}+U_{11}^{-1}+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot 2^{c-2}}\right)\\
&= \left(\frac{\frac{1}{\varepsilon}+\frac{U_{11}^{-1}}{2}}{\frac{1}{2^{n-1}\varepsilon}+\frac{U_{11}^{-1}}{2^{n-1}}+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot 2^{c-1-n}}\right)\theta
\end{aligned}
$$

will blow up at $\varepsilon = -\dfrac{1}{(U^{-1}L^{-1})_{11}} = -\dfrac{1}{(U_{11}^{-1}+\sum_{c=2}^{n}U_{1,c}^{-1}\cdot 2^{c-2})}$, which is usually a large value.

In particular, note that in the special case of the Wilkinson matrix, $U_{11}^{-1} = 1$ is the only nonzero entry in the first row, in which case $\dfrac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{11}}{(U^{-1}L^{-1})_{11}} \approx -p$ no longer holds, and instead,

$$
p_\varepsilon^{(1,1)} = \frac{\varepsilon+0.5}{\varepsilon+1}\cdot 2^{n-1} < 2^{n-1}
$$

In fact, similar computation would reveal that $p_\varepsilon^{(i,j)}$ approaches 2^{n-1} for all (i,j) distant from $(1,n-1)$, given a sufficiently small ε. In the case of a 25×25 Wilkinson matrix, for example, perturbing the $(1,1)$ entry with $\varepsilon = 1e-3$ yields a last pivot value of 1.68×10^7. On the other hand, the same perturbation applied to the $(1,n-1)$ entry results in a last pivot value of approximately 4.00×10^3, regardless of the matrix size n.

2.2 Patterns in Adjacent Entry Perturbations

Perturbing the $(1,1)$ entry results in the largest reduction of the last pivot value $p_\varepsilon^{(1,1)}$ of the random Higham^2 matrix in Figure 2.1. In fact, for a uniformly sampled Higham^2 matrix, $p_\varepsilon^{(1,1)}$ is usually the smallest among all $p_\varepsilon^{(i,j)}$ for $1 \leq i, j \leq n$. The following analysis suggests that for smaller indices i and j, perturbing the (i,j) entry yields a last pivot value that is approximately half as large as that resulting from perturbing the $(i+1,j)$ entry using the same ε.

Claim 2.2.1. *A uniformly sampled Higham^2 matrix with a sufficiently large size tends to have*

$$\frac{p_\varepsilon^{(i+1,j)}}{p_\varepsilon^{(i,j)}} \approx 2 \quad \textit{for small } i, j.$$

This assertion follows from the fact that

$$\begin{aligned}
\left(U^{-1}\boldsymbol{u}\ell^T L^{-1}\right)_{ji} &= \sum_{c=1}^{n} U^{-1}_{j,c} \cdot (-2^{c-i-1}p) = \sum_{c=j}^{n} U^{-1}_{j,c} \cdot (-2^{c-i-1}p) \\
\left(U^{-1}L^{-1}\right)_{ji} &= U^{-1}_{ji} + \sum_{c=1}^{n} U^{-1}_{j,c} L^{-1}_{c,i} = U^{-1}_{j,i} + \sum_{c=\max(j,i+1)}^{n} U^{-1}_{j,c} \cdot 2^{c-1-i}
\end{aligned}$$

Note that when $j > i$, $\dfrac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{ji}}{(U^{-1}L^{-1})_{ji}} = -p$,

When $j \leq i$, since U^{-1}_{ji} tends to have larger entries in magnitude at the upper right corner of U^{-1} (i.e., for i, j close to n) the latter terms tend to dominate and thus it is usually the case that $\dfrac{(U^{-1}\boldsymbol{u}\,\boldsymbol{\ell}^T L^{-1})_{ji}}{(U^{-1}L^{-1})_{ji}} \approx -p$.

As a consequence,

$$\begin{aligned}
p_\varepsilon^{(i,j)} &= p + \varepsilon \frac{\left(U^{-1}\boldsymbol{u}\ell^T L^{-1}\right)_{ji}}{1+\varepsilon\left(U^{-1}L^{-1}\right)_{ji}} \\
&= \frac{p + p\varepsilon\left(U^{-1}L^{-1}\right)_{ji} + \varepsilon\left(U^{-1}\boldsymbol{u}\ell^T L^{-1}\right)_{ji}}{1+\varepsilon\left(U^{-1}L^{-1}\right)_{ji}} \\
&\approx \frac{p}{1+\varepsilon\left(U^{-1}L^{-1}\right)_{ji}}.
\end{aligned}$$

Since

$$\left(U^{-1}L^{-1}\right)_{j,i+1} = U^{-1}_{j,i+1} + \sum_{c=\max(j,i+2)}^{n} U^{-1}_{j,c} \cdot 2^{c-i-2},$$

then

$$\frac{\left(U^{-1}L^{-1}\right)_{j,i}}{\left(U^{-1}L^{-1}\right)_{j,i+1}} = \frac{U^{-1}_{j,i} + U^{-1}_{j,i+1} + \sum_{c=\max(j,i+2)}^{n} U^{-1}_{j,c} \cdot \left(-2^{c-i-1}\right)}{U^{-1}_{j,i+1} + \sum_{c=\max(j,i+2)}^{n} U^{-1}_{j,c} \cdot 2^{c-i-2}} \approx 2 \quad \text{when } U^{-1}_{j,i} + U^{-1}_{j,i+1} \text{ are small.}$$

Thus we have the approximation

$$\begin{aligned}
\frac{p_\varepsilon^{(i+1,j)}}{p_\varepsilon^{(i,j)}} &\approx \frac{\frac{p}{1+\varepsilon\left(U^{-1}L^{-1}\right)_{j,i+1}}}{\frac{p}{1+\varepsilon\left(U^{-1}L^{-1}\right)_{j,i}}} \\
&\approx \frac{\left(U^{-1}L^{-1}\right)_{j,i}}{\left(U^{-1}L^{-1}\right)_{j,i+1}} \\
&\approx 2.
\end{aligned}$$

Although counterexamples such as the Wilkinson matrix, or matrices with small or cancelling entries near the upper right corner of U^{-1}, can demonstrate limitations to this approximation, empirical observations show that $\frac{p_\varepsilon^{(i+1,j)}}{p_\varepsilon^{(i,j)}} \approx 2$ holds in the vast majority of cases with uniformly sampled Higham [2] matrices. For matrices of size $n = 25$, this phenomenon is particularly evident at upper left indices (in particular, $i, j \leq 5$). These findings underscore the practical relevance of this approximation while also highlighting its dependency on specific matrix characteristics.

3

Error Propagation Analysis

In this chapter, we analyze how the initial perturbation, particularly at the top-left entry, propagates through the Gaussian elimination process. Theorem 3.1.1 illustrates how the error presented in the $(1, 1)$ entry of the initial Higham^2 matrix tends to double in magnitude with each step of the Gaussian elimination process. This provides an additional perspective on why perturbations at the top-left are more influential and more likely to disrupt the expected outcome of the Gaussian elimination process.

3.1 Propagation of Error at the Top Left Entry

Theorem 3.1.1. *Let A^* be an $n \times n$ Higham^2 matrix and A be a perturbed matrix of A^* with a perturbation ε at the (1, 1) entry. Let ${A^*}^{(k)}$ and $A^{(k)}$ denote the intermediate matrices after the k-th iteration*[1] *of precise Gaussian elimination on A^* and A, respectively. For any upper triangular entry (i, j) with $i \leq j$, the error is given by:*

$$A^{(k)}_{k+1,k+1:n} - {A^*}^{(k)}_{k+1,k+1:n} = \begin{cases} -\dfrac{\varepsilon}{A_{kk}} A_{k,k+1:n} & \textit{for} \quad k = 1 \\ -2\dfrac{(A^{(k-1)}_{kk} - {A^*}^{(k-1)}_{kk})}{A^{(k-1)}_{kk}} A^{(k-1)}_{k,k+1:n} + 2(A^{(k-1)}_{k,k+1:n} - {A^*}^{(k-1)}_{k,k+1:n}) & \textit{for} \quad k > 1 \end{cases}, \tag{3.1.1}$$

Theorem 3.1.1 demonstrates that, after the k-th iteration of the Gaussian elimination of a random Higham^2 matrix A without pivoting (i.e., after finishing processing the $(k+1)$-th row), the error in the $(k+1)$-th row arises from two sources: error in the multiplier $\left(-\dfrac{\varepsilon}{A_{kk}} A_{k,k+1:n} \text{ for } k = 1 \text{ and } -2\dfrac{(A^{(k-1)}_{kk} - {A^*}^{(k-1)}_{kk})}{A^{(k-1)}_{kk}} A^{(k-1)}_{k,k+1:n} \text{ for } k > 1\right)$ and the cumulative error from the previous row $\left(2(A^{(k-1)}_{k,k+1:n} - {A^*}^{(k-1)}_{k,k+1:n})\right)$. We discuss these two sources of error below.

3.2 Inexact Multiplier

For the Higham^2 matrix, theoretically, the factor at each step of the Gaussian elimination should be $\dfrac{A^{(k-1)}_{k+1,k}}{A^{(k-1)}_{kk}} = -1$. However, if there is an error in $A^{(k-1)}_{kk}$, the actual multiplier will deviate from -1. Assuming the actual multiplier is $-1 + \delta$, this results in each subsequent row receiving an additional contribution of $\delta A^{(k-1)}_{k,:}$. Consequently, for any row i below the k-th row, an error of $\delta A^{(k-1)}_{k,k+1:n}$ impacts the entries from $k+1$ to n.

[1] Using 1-based indexing, i.e., the first iteration uses the first row (without pivoting) as the pivot row to process all the rows below.

Note that when $k > 1$, when all entries in the k-th column are affected by an error η (due to the error from above it and affecting the k-th column in the preceeding iterations), the (k, k)-th entry has $A_{kk}^{(k-1)} = {A^*}_{kk}^{(k-1)} + \eta$ and all $(k+i, k)$-th entries below it has $A_{k+i,k}^{(k-1)} = -{A^*}_{kk}^{(k-1)} + \eta$. Therefore, the actual multiplier becomes

$$\frac{A_{k+i,k}^{(k-1)}}{A_{kk}^{(k-1)}} = \frac{-{A^*}_{kk}^{(k-1)} + \eta}{{A^*}_{kk}^{(k-1)} + \eta} = -1 + 2\frac{\eta}{{A^*}_{kk}^{(k-1)} + \eta} = -1 + \delta.$$

Note that $A_{kk}^{(k-1)} = {A^*}_{kk}^{(k-1)} + \eta$. This shows that the error in multiplier is

$$\delta = 2\frac{\eta}{{A^*}_{kk}^{(k-1)} + \eta} = -2\frac{(A_{kk}^{(k-1)} - {A^*}_{kk}^{(k-1)})}{A_{kk}^{(k-1)}}.$$

3.3 Error from the Previous Row

When the $(k-1)$-th pivot (and thus the $(k-1)$-th multiplier) contains error, the $(k-1)$-th iteration of the Gaussian elimination perturbs the k-th row not only in the pivot element (k, k), but also in all the remaining elements $A_{k,k+1:n}^{(k-1)}$.

Let the error in the remaining elements of the k-th row be u, where $u = A_{k,k+1:n}^{(k-1)} - {A^*}_{k,k+1:n}^{(k-1)}$. Note that this error vector u is added to the corresponding part (from the $(k+1)$-th to the n-th entries) of every row below the $(k-1)$-th row. That is, $A_{k+i,k+1:n}^{(k-1)} - {A^*}_{k+i,k+1:n}^{(k-1)} = u \ \forall i$. Due to the -1 multiplier (i.e., processing each row involves adding this row to all the subsequent rows below it), when the k-th row is processed in the k-th iteration, the error in the $(k+1)$-th to the n-th entries now becomes $2\,u$. Note that the factor of 2 here is precise because the error in multiplier has been accounted for in the previous section precisely.

3.4 Generalization to Perturbations at Other Entries

While the above discussion only assumes an initial error ε in the (1, 1) entry of the matrix, the analysis can be generalized to an initial error in any entry:

- For any diagonal (j, j) or upper diagonal entry (i, j) with an error, an inexact multiplier arises at the j-th row, from which error propagation begins.
- For any lower diagonal entry (i, j) with an error, both an inexact multiplier and errors from previous rows arise starting at the i-th row, leading to subsequent error propagation.

3.5 Implications

In practice, for a general Higham[2] matrix, the numerical results show that the magnitudes of the two error sources are comparable, with neither source consistently predominating.

As Equation (3.1.1) suggests, for both sources of error, since the error from the initial iteration will propagate through all the subsequent iterations, their magnitude tends to double with each iteration of Gaussian elimination. Although there is no guarantee for the sign of the errors and there is potential for the errors to cancel each other out, they generally show a strong tendency to grow exponentially. For example, when $n \geq 30$, even machine-epsilon level of error could blow up, so the expected growth of 2^{n-1} is no longer observed. This also illustrates why perturbations closer to the top left entry tend to have a larger effect.

Nevertheless, we can still manually construct cases for which $\lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} p_\varepsilon^{(1,1)} = 0$ so that this error ε applied at (1, 1) does not have any effect on the last pivot. For example, with $n = 3$, the expression

$$t_{13} = \frac{4t_{12}t_{23} - 2t_{12}t_{33} + t_{22}t_{33}}{4t_{22}}$$

demonstrates such a scenario. Similarly, for $n = 4$,

$$t_{14} = \frac{8t_{12}t_{24}t_{33} + 8t_{13}t_{22}t_{34} - 8t_{12}t_{23}t_{34} - 4t_{13}t_{22}t_{44} + 4t_{12}t_{23}t_{44} - 2t_{12}t_{33}t_{44} + t_{22}t_{33}t_{44}}{8t_{22}t_{33}}$$

provides another example. Numerical computations confirm that in these setups, the two sources of error precisely offset one another in the final iteration, and consequently the growth remains exactly 2^{n-1}.

4

Impact of Top Left Entry Perturbation

Not all directions of perturbation have equal effect. Given the observation that perturbing the $(1, 1)$ entry tends to be the most effective for a general Higham ^{2} matrix, Figure 4.1 demonstrates the effect on the last pivot of applying to a random Higham^2 matrix A a perturbation of P_1, where

$$P_1[i, j] = \begin{cases} N(0, 1) & \text{if } i = 1 \text{ and } j = 1 \\ 0 & \text{otherwise} \end{cases}$$

(only perturbing the $(1,1)$ entry) versus P_2, where

$$P_2[i,j] = \begin{cases} 0 & \text{if } i = 1 \text{ and } j = 1 \\ \frac{1}{\sqrt{n}} \cdot N(0,1) & \text{otherwise} \end{cases}$$

(perturbing all the remaining entries except the $(1,1)$ entry).

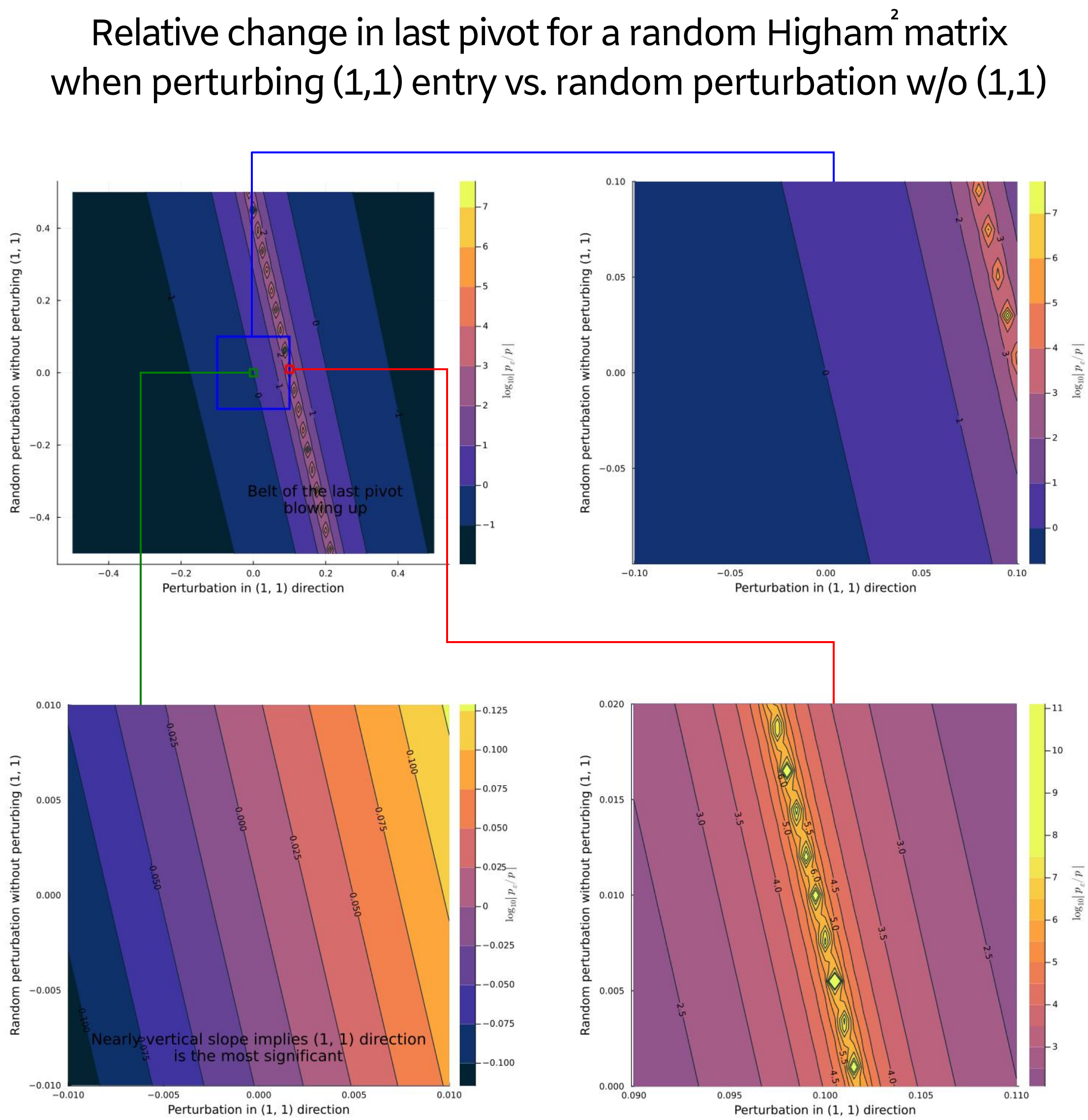

Figure 4.1: Contour plots of the log ratio of the perturbed and the original last pivot $\log \left|\frac{p_\varepsilon}{p}\right|$ of a random Higham^2 matrix.

p is the last pivot of a random Higham^2 matrix A, and p_ε is the growth of the perturbed matrix $A + \varepsilon_1 P_1 + \varepsilon_2 P_2$, where P_1 is a zero matrix with normal random value at the $(1, 1)$ entry, and P_2 is a random matrix randn(n, n) / $\sqrt{n}$ but the entry $(1, 1)$ is set to zero.
The upper left plot visualizes the effect of $\varepsilon_1, \varepsilon_2 \in [-0.5, 0.5]$.
The upper right plot visualizes the effect of $\varepsilon_1, \varepsilon_2 \in [-0.1, 0.1]$ (the area of the green box in the upper left plot).
The lower left plot visualizes the effect of $\varepsilon_1, \varepsilon_2 \in [-0.01, 0.01]$ (the area of the blue box in the upper left plot).
The lower right plot visualizes an area of blow-up, where $\varepsilon_1 \in [0.09, 0.11]$ and $\varepsilon_2 \in [0, 0.02]$ (the area of the red box in the upper left plot).

4.1 Predominant Influence of Upper Left Corner Perturbations

Figure 4.1 predominantly displays contour lines as tilted parallel lines across the entire range of $\varepsilon_1, \varepsilon_2$, which suggests that the effects of the two perturbations are largely independent and separable across most regions. The steep slope suggests that perturbations in the P_1 direction have a more pronounced effect on the last pivot compared to the P_2 direction. In fact, despite perturbations affecting almost all entries, the major contributors to the reduction of the last pivot in P_2 are still the entries close to $(1, 1)$, such as $(2, 1)$ and $(1, 2)$. Perturbations in entries in the upper right corner exhibit negligible effects on the last pivot.

4.2 Blow-Up Phenomena

Figure 4.1 also displays a distinct band of blow-up region where the last pivot value spikes, and the contours in this area are more intricate. This occurs at a specific ratio of ε_1 and ε_2, depending on the random values within P_1 and P_2. The presence of this blow-up band can be attributed to the fact that the last pivot is expressible in terms of ε as $p_\varepsilon = \dfrac{a\varepsilon + b}{c\varepsilon + d}$. According to Equation 1.1.1, the blow-up occurs when the denominator $1 + \varepsilon(LU)^{-1}_{ji} = 0$, i.e., when $\varepsilon = -\frac{1}{(LU)^{-1}_{ji}}$, though such value is typically not achieved in practice. As the value of $(\varepsilon_1, \varepsilon_2)$ moves farther from the band of blow-up region, the magnitude of the last pivot diminishes, although the rate of reduction tends to decelerate.

5

Perturbation for Growth Reduction

Since perturbing the $(1, 1)$ entry effectively reduces the growth of most Higham^2 matrices, and special cases like the Wilkinson matrix can be handled by perturbing the $(1, n-1)$ entry, a comprehensive approach for consistently reducing the last pivot value and the overall growth is to perturb all entries in the first row. Each entry affects the last pivot of some submatrix within the full matrix, making this a reliable method to control the growth of the entire class of the Higham^2 matrices.

5.1 Strategic Perturbation on the First Row

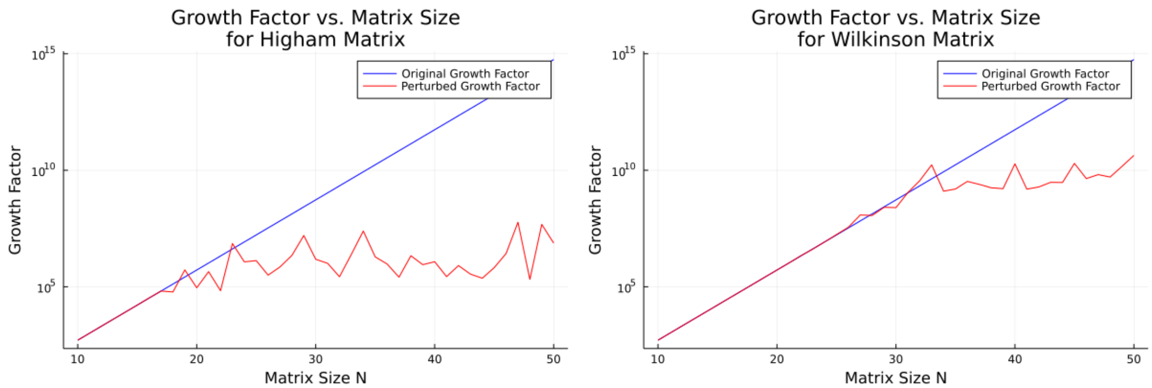

Figure 5.1: Plot of the growth factor of A and the perturbed matrix $A + P$ versus the matrix size n. A is a uniformly sampled Higham^2 matrix in the left plot and a Wilkinson matrix in the right plot. The perturbation $P = \varepsilon \boldsymbol{e}_1 \cdot \boldsymbol{v}^\top$, with $v_i \sim N(0,1)$ and $\varepsilon = 1e - 10$, applies a normal random perturbation across the first row. While the unperturbed matrix growth follows the form of 2^{n-1}, the growth of the perturbed Higham^2 and Wilkinson matrices plateau at approximately $1e6$ and $1e10$ for matrix sizes around 18 and 30, respectively.

The effect of this perturbation strategy is illustrated in Figure 5.1. When we apply a row of random normal numbers with a standard deviation of $1e - 8$, the growth of a random Higham^2 matrix and the Wilkinson matrix levels off at around $1e6$ and $1e10$ at matrix sizes of approximately 18 and 30, respectively. This indicates that Higham^2 matrices lie on a ”ridge” in the matrix space, where even tiny random perturbations can significantly affect their behavior in Gaussian elimination and make them drop from the theoretical high growth.

5.2 Round-off Errors in Practical Computations

In fact, in practical computations with floating-point arithmetics, the error presented in the initial floating-point representation of the Higham^2 matrix is sufficient to distort the Gaussian elimination calculation, given that the matrix is sufficiently large size (e.g., $n = 40$), let alone the errors arising from the subsequent floating-point computations.

As discussed in Chapter 3, due to the way Higham^2 matrices are constructed, each iteration of the Gaussian elimination process involves directly adding the adjacent rows. This iterative process of adding neighboring rows can make any errors initially present at the top of the matrix not only persist but double with each row processing, leading to exponential error propagation. Thus, it is interesting to observe that the very structural characteristics that make Higham^2 matrices achieve such high theoretical growth factors also limit their growth in practical computations due to round-off errors. In practice, even the presence of the machine epsilon itself would prevent consistently achieving any growth greater than $1e11$ in double-precision floating-point computations, making the theoretical growth unachievable for random Higham^2 matrices larger than 40×40. These observations show how important and effective it is to use small perturbations in the initial entries to guarantee matrix stability under exact arithmetics (in the absence of floating-point errors), especially in applications that require high numerical precision.

6
Conclusion

In this thesis, we have analyzed the behaviors of Higham^2 matrices under small perturbations through theoretical calculations and empirical observations to examine the stability of Gaussian elimination under extreme scenarios. Gaussian elimination is a fundamental algorithm in linear algebra. However, it still can suffer from errors, more so when considering large-scale ill-conditioned problems. We thus focus on the class of Higham^2 matrices, including the special case of Wilkinson matrices, to understand the effect of entrywise perturbations on the Gaussian elimination process. The focus of our analysis on Higham^2 matrices allows us to extract specific numerical properties from these particular matrix structures and algorithms.

Theorem 1.2.1 suggests that Higham^2 matrices, when subjected to perturbations, show complex but predictable behaviors. Even though there are special cases such as Wilkinson matrices, perturbing the top left entries tends to be the most effective at reducing the growth of a general Higham^2 matrix. The upper bound of $|p_\varepsilon^{(1,n-1)}|$ also suggests that strategically perturbing the first row can consistently reduce the growth, thus enhancing the robustness of numerical methods against instabilities caused by such adverse matrix properties. Section 5.1 demonstrates the effectiveness and reliability of this method for controlling the growth factor of an arbitrary Higham^2 matrix. We also point out in Section 5.2 that while Higham^2 matrices theoretically pose a significant challenge in numerical operations like solving $Ax = b$, their inherent sensitivity to even minor perturbations often neutralizes this potential threat. Our findings elucidate how the very propensity for high growth in these matrices makes them unlikely to manifest such extremes in practical computational scenarios, where floating-point errors are almost inevitable.

We have also provided an analysis of the error propagation in the Gaussian elimination of Higham^2 matrices in Chapter 3. Theorem 3.1.1 demonstrates how errors, especially at the top-left entry, propagate and tend to double in magnitude with each step. This analysis provides an additional understanding of why perturbations at the top-left are particularly disruptive. We discuss the two primary sources of error—error in the multiplier and cumulative error from previous rows—and how these errors compound through the elimination process.

Finally, the numerical experiments that we have conducted provide further evidence supporting our theoretical analysis. The heatmaps and contour plots visualize the behavior of Higham^2 matrices under perturbations, aligning with the theoretical predictions that we have established in Chapter 1 and Chapter 2. These results illustrate the practical impact of entrywise perturbations on matrix stability.

References

[1] Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1), 65–98.

[2] Bisain, A., Edelman, A., & Urschel, J. (2023). A new upper bound for the growth factor in gaussian elimination with complete pivoting. *arXiv preprint arXiv:2312.00994*.

[3] Edelman, A. & Urschel, J. (2024). Some new results on the maximum growth factor in gaussian elimination. *SIAM Journal on Matrix Analysis and Applications*, 45(2), 967–991.

[4] Foster, L. V. (1994). Gaussian elimination with partial pivoting can fail in practice. *SIAM Journal on Matrix Analysis and Applications*, 15(4), 1354–1362.

[5] Golub, G. H. & Van Loan, C. F. (2013). *Matrix computations*. JHU press.

[6] Gould, N. (1991). On growth in Gaussian elimination with complete pivoting. *SIAM Journal on Matrix Analysis and Applications*, 12(2), 354–361.

[7] Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., & Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362.

[8] Higham, D. J., Higham, N. J., & Pranesh, S. (2021). Random matrices generating large growth in lu factorization with pivoting. *SIAM Journal on Matrix Analysis and Applications*, 42(1), 185–201.

[9] Higham, N. J. & Higham, D. J. (1989). Large growth factors in Gaussian elimination with pivoting. *SIAM Journal on Matrix Analysis and Applications*, 10(2), 155–164.

[10] Huang, H. & Tikhomirov, K. (2022). Average-case analysis of the Gaussian elimination with partial pivoting. *arXiv preprint arXiv:2206.01726*.

[11] R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

[12] Sankar, A. (2004). *Smoothed analysis of Gaussian elimination*. PhD thesis, Massachusetts Institute of Technology.

[13] Sankar, A., Spielman, D. A., & Teng, S.-H. (2006). Smoothed analysis of the condition numbers and growth factors of matrices. *SIAM Journal on Matrix Analysis and Applications*, 28(2), 446–476.

[14] Spielman, D. A. & Teng, S.-H. (2004). Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM (JACM)*, 51(3), 385–463.

[15] The MathWorks Inc. (2022). Matlab version: 9.13.0 (r2022b).

[16] Trefethen, L. N. & Bau, D. (2022). *Numerical linear algebra*, volume 181. Siam.

[17] Trefethen, L. N. & Schreiber, R. S. (1990). Average-case stability of Gaussian elimination. *SIAM Journal on Matrix Analysis and Applications*, 11(3), 335–360.

[18] Wilkinson, J. (1965). *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford.

[19] Wolfram Research, Inc. (2024). Mathematica, Version 14.0. Champaign, IL.