# CONDITION NUMBERS IN TENSOR DECOMPOSITIONS

NICK VANNIEUWENHOVEN

Abstract. I report on the progress made in the area of condition numbers in mutlilinear algebra during the semester *"AGATES: Algebraic Geometry with Applications to TEnsors and Secants."* These notes report on very preliminary research result and may hence contain errors or inaccuracies.

## 1. Introduction

Rice [13] gave a general definition of the condition number of $\phi : \mathcal{I} \to \mathcal{O}$ that applies to *metric* spaces $(\mathcal{I}, \mathrm{dist}_\mathcal{I})$ and $(\mathcal{O}, \mathrm{dist}_\mathcal{O})$, namely

$$\kappa[\phi](x) := \lim_{\epsilon \to 0} \sup_{\mathrm{dist}_\mathcal{I}(x,x') \le \epsilon} \frac{\mathrm{dist}_\mathcal{O}(\phi(x), \phi(x'))}{\mathrm{dist}_\mathcal{I}(x, x')}. \tag{1}$$

For example, a *Riemannian manifold*, a smooth manifold equipped with a *Riemannian metric* (i.e., a smooth tensor field that restricts to an inner product on each tangent space), is a metric space [11]. In this case, Rice [13] states that the above condition number in eq. (1) satisfies

$$\kappa[\phi](x) = \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}} := \sup_{t_x \in \mathrm{T}_x\mathcal{I}} \frac{\|(\mathrm{d}_x\phi)(t_x)\|_{\mathcal{O},\phi(x)}}{\|t_x\|_{\mathcal{I},x}}, \tag{2}$$

where $\mathrm{d}_x\phi$ is the derivative of the smooth map $\phi : \mathcal{I} \to \mathcal{O}$.

It is well known to the experts that symmetry of the metrics $\mathrm{dist}_\mathcal{I}$ and $\mathrm{dist}_\mathcal{O}$ is not required in eq. (1) because of the natural asymmetry between the original point $x$ (respectively $\phi(x)$) and the perturbed point $x'$ (respectively $\phi(x')$) [2, Chapter 12]. Indeed, the condition number is defined specifically at the designated (unperturbed) point $x$. When $\phi : \mathcal{I} \to \mathcal{O}$ is a map between smooth manifolds, it is therefore sufficient to equip $\mathcal{I}$ and $\mathcal{O}$ with *smooth Finsler metrics.* A Finsler metric is a smooth tensor field on the tangent bundle minus the zero section that restricts to an asymmetric norm (a function $\nu$ that satisfies all axioms of a norm except for $\nu(x) = \nu(-x)$) on each tangent space. I show that the following generalization[1] of Rice's condition number applies in the setting of smooth Finsler manifolds, i.e., smooth manifolds equipped with a smooth Finsler metric.

**Theorem 1.** *Let $(\mathcal{I}, \|\cdot\|_\mathcal{I})$ and $(\mathcal{O}, \|\cdot\|_\mathcal{O})$ be smooth Finsler manifolds with induced local quasimetrics $\mathrm{dist}_\mathcal{I}$ and $\mathrm{dist}_\mathcal{O}$ respectively. Then, the condition number of a smooth map $\phi : \mathcal{I} \to \mathcal{O}$ at $x \in \mathcal{I}$ is*

$$\kappa[\phi](x) := \lim_{\epsilon \to 0} \sup_{\mathrm{dist}_\mathcal{I}(x,x') \le \epsilon} \frac{\mathrm{dist}_\mathcal{O}(\phi(x), \phi(x'))}{\mathrm{dist}_\mathcal{I}(x, x')} = \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}.$$

This work is partially supported by the Thematic Research Programme "Tensors: geometry, complexity and quantum entanglement," University of Warsaw, Excellence Initiative – Research University and the Simons Foundation Award No. 663281 granted to the Institute of Mathematics of the Polish Academy of Sciences for the years 2021-2023.

[1] Riemannian manifolds are automatically Finsler manifolds when the norm induced from the Riemannian metric is the Finsler metric.

Note that the new insight, relative to [2, Chapter 12], is that this theorem shows that Rice's condition number, defined as a limit superior of a fraction of (quasi)metrics on the manifold, equals the spectral norm of the derivative $\|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}$, whereas [2] takes the latter as a definition. As an immediate consequence the following result is obtained.

**Corollary 1.** *The following bound is first-order sharp:*

$$\mathrm{dist}_{\mathcal{O}}(\phi(x),\phi(x')) \le \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}} \cdot \mathrm{dist}_{\mathcal{I}}(x,x') + o(\mathrm{dist}_{\mathcal{I}}(x,x')).$$

The above upper bound is first-order sharp in the sense that it is an equality when $x$ and $x'$ are connected by a minimizing geodesic whose tangent vector at $x$ is the right singular vector of $\mathrm{d}_x\phi$ that corresponds to the largest singular value, namely $\|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}$.

Recall that in numerical analysis condition numbers are often categorized as either "absolute" or "relative." This refers to qualitative aspects of the (quasi)metrics used in the definition of the condition number in eq. (1) and theorem 1. An *absolute condition number* is intended to measure errors in an absolute sense, much like the metric $\mathrm{dist}_{\mathbb{R}}(x,y) = |x-y|$ does not take into account the magnitude of $x$ and $y$. By contrast, a *relative condition number* is intended to refer to a condition number with respect to (quasi)metrics that measure distance in some relative sense, much like Pryce's relative error metric $\mathrm{dist}_{\mathbb{R}}^{\mathrm{Pryce}}(x,y) = \left|\log\frac{x}{y}\right|$ from [12]. I will use $\kappa$ to denote condition numbers that morally measure distances in an absolute sense, and $\mu$ to denote condition numbers that morally measure distances in a relative sense.

By applying the foregoing theorem, the following interesting result about the absolute and relative condition number of matrix multiplication is obtained. I could not locate this precise statement in the literature, which may hence be novel.

**Theorem 2.** *Let*

$$\cdot : \underbrace{\mathbb{R}_*^{m\times n} \times \mathbb{R}_*^{n\times p}}_{\mathcal{I}} \to \underbrace{\mathbb{R}^{m\times p}}_{\mathcal{O}},\ (A,B)\mapsto AB$$

*denote the multiplication of nonzero matrices. Then, the absolute condition number of matrix multiplication with respect to the product Frobenius norm on $\mathcal{I}$ and the Frobenius norm on $\mathcal{O}$ is*

$$\kappa[\cdot](A,B) = \sqrt{\|A\|_2^2 + \|B\|_2^2}.$$

*The relative condition number of matrix multiplication with respect to the Finsler metrics*

$$\|(\dot{A},\dot{B})\|_{\mathcal{I},(A,B)} = \sqrt{\frac{\|\dot{A}\|_F^2}{\|A\|_2^2} + \frac{\|\dot{B}\|_F^2}{\|B\|_2^2}} \quad \text{and} \quad \|\dot{X}\|_{\mathcal{O},X} = \frac{\|\dot{X}\|_F}{\|X\|_2}$$

*on $\mathcal{I}$ and $\mathcal{O}$, respectively, is*

$$\mu[\cdot](A,B) = \frac{\|A\|_2\|B\|_2}{\|AB\|_2}.$$

**Notation.** The vector norm $\|\cdot\|$ without subscripts refers to the Euclidean norm. The length-$n$ vector which is zero everywhere except for a 1 in position $i$ is denoted by $e_i \in \mathbb{R}^n$.

**Outline.** In the next section, I prove theorem 1. I present some illustrative calculations of well-known condition numbers in linear algebra in section 3. Thereafter, in section 4, theorem 2 is proved. Finally, I compute the condition number of applying an arbitrary multilinear map in section 5.

## 2. The condition number of maps between Finsler manifolds

Observe that by definition, the spectral norm from eq. (2) is consistent in the sense that $\|Ax\|_{\mathcal{O}} \le \|A\|_{\mathcal{I}\to\mathcal{O}}\|x\|_{\mathcal{I}}$.

The length of a smooth curve $\gamma(t) \subset \mathcal{I}$ in a Finsler manifold $(\mathcal{I}, \|\cdot\|_{\mathcal{I}})$ from 0 to $s$ is defined as

$$L[\gamma(s)] := \int_0^s \|\gamma'(t)\|_{\mathcal{I},\gamma(t)} \mathrm{d}t,$$

where $\gamma'(t)$ is the derivative of $\gamma$ at $t$. With this definition in place, a Finsler manifold can be turned into a quasimetric space, by defining the distance between $x, y \in \mathcal{I}$ as the infumum of the lengths of piecewise smooth curves connecting $x$ and $y$.

We are now ready to prove theorem 1.

*Proof of theorem 1.* We show that $\kappa[\phi](x) \le \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}$. Assume that $\epsilon > 0$ is sufficiently small. Let $\gamma$ be any geodesic connecting $x$ and $x'$. We assume this geodesic is parameterized such that $\gamma(0) = x$ and $\gamma(\epsilon) = x'$. Then, we have

$$\begin{aligned}
\kappa_\epsilon[\phi](x) &:= \sup_{\mathrm{dist}_{\mathcal{I}}(x,x')=\epsilon} \frac{\mathrm{dist}_{\mathcal{O}}(\phi(x), \phi(x'))}{\mathrm{dist}_{\mathcal{I}}(x, x')} \\
&\le \sup_{\mathrm{dist}_{\mathcal{I}}(x,x')=\epsilon} \frac{1}{\epsilon} L[(\phi \circ \gamma)(\epsilon)] \\
&= \sup_{\mathrm{dist}_{\mathcal{I}}(x,x')=\epsilon} \frac{1}{\epsilon} \int_0^\epsilon \|(\mathrm{d}_{\gamma'(t)}\phi)(\gamma'(t))\|_{\mathcal{O},\gamma(t)} \mathrm{d}t \\
&\le \sup_{\mathrm{dist}_{\mathcal{I}}(x,x')=\epsilon} \frac{1}{\epsilon} \left( \sup_{0\le t\le\epsilon} \|\mathrm{d}_{\gamma(t)}\phi\|_{\mathcal{I}\to\mathcal{O}} \right) \int_0^\epsilon \|\gamma'(t)\|_{\mathcal{O},\gamma(t)} \mathrm{d}t \\
&= \sup_{\mathrm{dist}_{\mathcal{I}}(x,x')=\epsilon} \frac{L[\gamma(\epsilon)]}{\epsilon} \sup_{0\le t\le\epsilon} \|\mathrm{d}_{\gamma(t)}\phi\|_{\mathcal{I}\to\mathcal{O}}.
\end{aligned}$$

where the first inequality used that the distance between $\phi(x)$ and $\phi(x')$ is defined as the infimum over all smooth curves. Since $\gamma$ *is* a geodesic, $L[\gamma(\epsilon)] = \epsilon$. Hence, taking limits as $\epsilon \to 0$, we obtain $\kappa[\phi](x) \le \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}$.

Conversely, we show that $\kappa[\phi](x) \ge \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}$. Consider a minimizing geodesic $\gamma$ of length $\epsilon$ connecting $x \in \mathcal{I}$ and $x' \in \mathcal{I}$ such that $\|\gamma'(0)\|_{\mathcal{I}} = \epsilon$ and $\|(\mathrm{d}_x\phi)\gamma'(0)\|_{\mathcal{O}} = \epsilon\|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}}$.[2] Let $v = (\mathrm{d}_x\phi)\gamma'(0)$ and consider the two smooth curves

$$\tau_1 : t \mapsto tv \quad \text{and} \quad \tau_2 : t \mapsto \exp_{\phi(x)}^{-1}(\phi(\gamma(t))).$$

Note that the second curve is well defined because the exponential map provides a diffeomorphism between a neighborhood of 0 in $\mathrm{T}_{\phi(x)}\mathcal{O}$ (after using the canonical isomorphism between $\mathrm{T}_0\mathrm{T}_{\phi(x)}\mathcal{O}$ and $\mathrm{T}_{\phi(x)}\mathcal{O}$) and a neighborhood of $\phi(x)$ in $\mathcal{O}$ [11, Chapter 5, Proposition 18]. These two curves match to first order. Indeed, $\tau_1(0) = \tau_2(0) = \phi(x)$, $\tau_1'(0) = v$, and

$$\tau_2'(0) = (\mathrm{d}_{\phi(\gamma(0))} \exp_{\phi(x)}^{-1})(\mathrm{d}_{\gamma(0)}\phi)\gamma'(0) = (\mathrm{d}_{\phi(x)} \exp_{\phi(x)})^{-1} v = v,$$

where the second equality is by the inverse function theorem (and that $\exp_p$ is a diffeomorphism), and the third used the fact that $\mathrm{d}_0 \exp_p = \mathrm{Id}_{\mathrm{T}_p\mathcal{O}}$; see, e.g., [11, Chapter 5, proof of Proposition 18]. Taking Taylor series of $\tau_1 \subset \mathrm{T}_{\phi(x)}\mathcal{O} \simeq \mathbb{R}^{\dim N}$ and $\tau_2 \subset \mathrm{T}_{\phi(x)}\mathcal{O}$, we find $\tau_1(t) = 0 + tv$ and $\tau_2(t) = 0 + tv + \mathcal{O}(t^2)$. Hence, $\|\tau_1(\epsilon) - \tau_2(\epsilon)\|_{\mathcal{O}} = \mathcal{O}(\epsilon^2)$. By construction, we have

$$\|\tau_2(\epsilon)\|_{\mathcal{O}} = \| \exp_{\phi(x)}^{-1}(\phi(x'))\|_{\mathcal{O}} = \mathrm{dist}_{\mathcal{O}}(\phi(x), \phi(x')).$$

[2] The map $\phi$ between $(\mathcal{I}, g)$ and $(\mathcal{O}, h)$ is in general not a (local) isometry, so the curve $\phi(\gamma(t))$ is not usually a geodesic.

Exploiting the triangle inequality in $\mathrm{T}_{\phi(x)}\mathcal{O}$, we find that

$$\mathrm{dist}_{\mathcal{O}}(\phi(x), \phi(x')) = \|\tau_2(\epsilon)\|_{\mathcal{O}} \geq \|\tau_1(\epsilon)\|_{\mathcal{O}} - \mathcal{O}(\epsilon^2) = \epsilon\|v\|_{\mathcal{O}} - \mathcal{O}(\epsilon^2).$$

As a result, we obtain the desired bound:

$$\kappa_\epsilon[\phi](x) = \sup_{\mathrm{dist}_{\mathcal{I}}(x,x')=\epsilon} \frac{\mathrm{dist}_{\mathcal{O}}(\phi(x), \phi(x'))}{\mathrm{dist}_{\mathcal{I}}(x, x')} \geq \frac{1}{\epsilon}(\epsilon\|v\|_{\mathcal{O}} - \mathcal{O}(\epsilon^2))$$
$$= \|\mathrm{d}_x\phi\|_{\mathcal{I}\to\mathcal{O}} - \mathcal{O}(\epsilon).$$

Taking the limit $\epsilon \to 0$ concludes the proof. □

## 3. Condition in linear algebra

I give some basic, known examples of condition numbers of standard computational problems in numerical linear algebra, namely applying a linear map, inverting a matrix, and computing the kernel of a matrix. The examples in this section are intended to be of educational value, showing how condition numbers can be computed in practice.

The condition numbers of applying a linear map and computing the matrix inverse are foundational results in numerical analysis; see, e.g., [7, 3]. The condition number of computing the kernel of a matrix is probably not so well known because the output of this computational problem lives on the manifold of linear subspaces; nevertheless it is computed in detail by Bürgisser and Cucker [3].

3.1. **Applying a linear map.** Since the derivative of a linear map $A : \mathcal{I} \to \mathcal{O}$ between normed vector spaces $(\mathcal{I}, \|\cdot\|_{\mathcal{I}})$ and $(\mathcal{O}, \|\cdot\|_{\mathcal{O}})$ is the map itself, it follows immediately that its absolute condition number is the spectral norm of $A$ in the chosen norms. In formulas:

$$\kappa[A](x) = \|A\|_{\mathcal{I}\to\mathcal{O}}.$$

3.2. **Inverting a matrix.** Next, I consider the problem of computing the inverse of an invertible matrix $A \in \mathrm{GL}(n, \mathbb{R})$. This example is also presented by Higham [7, Theorem 6.4] and Bürgisser and Cucker [3, Theorem 1.5] in a more general form.

Note that the set of invertible matrices is not linear, but rather it is a special smooth manifold that admits a group structure under matrix multiplication (a Lie group). Hence, the map whose condition number we want to determine is

$$\cdot^{-1} : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R}), \quad A \mapsto A^{-1}.$$

The derivative of this map is easily obtained from the characterization $AA^{-1} = I$. Indeed, by deriving this equation on both sides we find an equality of linear maps

$$\dot{A} \mapsto \dot{A}A^{-1} + A((\mathrm{d}_A\cdot^{-1})(\dot{A})) = 0.$$

Consequently, for every $\dot{A} \in \mathrm{T}_A\mathrm{GL}(n, \mathbb{R})$, we have

$$(\mathrm{d}_A\cdot^{-1})(\dot{A}) = -A^{-1}\dot{A}A^{-1}.$$

As the derivative $\mathrm{d}_A\cdot^{-1} : \mathrm{T}_A\mathrm{GL}(n, \mathbb{R}) \to \mathrm{T}_A\mathrm{GL}(n, \mathbb{R})$ is a linear map, the above expression completely defines it.

If we take the standard basis matrices $e_i e_j^T$ in lexicographic order, then it is a standard result [8, Chapter 4] that the matrix of $X \mapsto PXQ$ can be represented as $P \otimes Q^T$, where $\otimes$ is the *tensor product* (of the linear maps represented in bases by $P$ and $Q^T$; see [6]).[3] Hence, if we choose the Frobenius norm as metric on the

[3]This tensor product of linear maps or the matrices representing them is also known as the Kronecker product.

input and output space, then we find

$$\kappa[\cdot^{-1}](A) = \sup_{\Delta \in \mathbb{R}^{n\times n}} \frac{\| - A^{-1}\Delta A^{-1}\|_F}{\|\Delta\|_F} = \sup_{\Delta \in \mathbb{R}^{n\times n}} \frac{\|(A^{-1} \otimes A^{-T})\mathrm{vec}(\Delta)\|_2}{\|\Delta\|_F}.$$

Since $(\mathbb{R}^{n\times n}, \|\cdot\|_F)$ is isometric to $(\mathbb{R}^{n^2}, \|\cdot\|_2)$, we can conclude that the absolute condition number of matrix inversion is given by the spectral norm of $A^{-1} \otimes A^{-T}$, which equals

$$\kappa[\cdot^{-1}](A) = \|A^{-1} \otimes A^{-T}\|_2 = \|A^{-1}\|_2^2.$$

The above condition number does not yet correspond to the well known "matrix condition number" $\kappa(A) = \|A\|_2\|A^{-1}\|_2$. However, by choosing the Finsler metrics

$$\|\dot{A}\|_{\mathcal{I},A} = \frac{\|\dot{A}\|_F}{\nu_1(A)} \quad \text{and} \quad \|\dot{A}\|_{\mathcal{O},A} = \frac{\|\dot{A}\|_F}{\nu_2(A)},$$

where $\nu_1$ and $\nu_2$ are any norms on $\mathbb{R}^{n\times n}$, we obtain the relative condition number

$$\mu[\cdot^{-1}](A) = \|A^{-1}\|_2^2 \frac{\nu_1(A)}{\nu_2(A^{-1})}.$$

In particular, choosing $\nu_1 = \nu_2 = \|\cdot\|_2$ results in the usual matrix condition number $\mu[\cdot^{-1}](A) = \|A\|_2\|A^{-1}\|_2 = \kappa(A)$. Hence, the matrix condition number is a relative condition number of the computational problem of inverting a matrix.

3.3. **Computing the kernel.** Consider the problem of computing the kernel of a linear map $A : V \to W$ between the $m$-dimensional real vector space $V$ and $n$-dimensional real vector space $W$. The kernel is the linear subspace of $V$ that is mapped to 0 under $A$. This means that the output of the computational problem "Compute the kernel of $A$" necessarily lives in the set of $(m-n)$-dimensional linear subspaces of $V$. The set of $k$-dimensional subspace of $V$ is called the *Grassmannian* (of $k$-dimensional subspaces of $V$):

$$\mathrm{Gr}(k, V) = \{A \text{ is a linear subspace of } V \mid \dim A = k\}.$$

The Grassmannian has many equivalent formulations. For example, it can be realized concretely as the set of rank-$k$ orthogonal projectors:

$$\mathrm{Gr}(k, V) = \{QQ^T \mid Q^T Q = I_k\} \subset \mathbb{R}^{m\times m},$$

where $V$ is assumed to be an $m$-dimensional real vector space. It is clear from this formulation that $\mathrm{Gr}(k, V)$ is not a linear space, as a sum of two rank-$k$ matrices is in general a matrix of rank $\min\{2k, m\}$. Instead, it is known that the Grassmannian is a smooth manifold [9].

Assume that $A$ is expressed with respect to orthogonal bases as the $n \times m$ matrix $A$. We can use theorem 1 to compute the condition number of the concrete problem

$$\ker : \mathbb{R}_\star^{n\times m} \to \mathrm{Gr}(m-n, \mathbb{R}^m), \; A \mapsto UU^T,$$

where $\mathbb{R}_\star^{n\times m}$ denotes the manifold of matrices with full (row) rank and $UU^T$ is (a projector onto) the kernel of $A$, when viewing Gr as a submanifold of $\mathbb{R}^{m\times m}$.

Next, I present one way to perform this computation. An alternative approach of similar complexity to arrive at the same result is given in [3, Section 14.3.2].

Let $U \in \mathrm{St}(m, m-n)$ be an $m \times (m-n)$ matrix with orthonormal columns. The span of $U$ is the $(m-n)$-dimensional kernel of $A$ if and only if

$$AUU^T = 0.$$

Taking the derivative of this equation in $A$ and $U$, we find jointly

$$0 = A(\dot{U}U^T + U\dot{U}^T) + \dot{A}UU^T = A\dot{U}U^T + \dot{A}UU^T = A\dot{U} + \dot{A}U, \tag{3}$$

where $\dot{A} \in \mathrm{T}_A\mathbb{R}^{n\times m}$ and $\dot{U} \in \mathrm{T}_U\mathrm{St}(m, m-n)$. Note that $U$ depends on $A$, but this is not indicated in the notation.

It is well known [5] that the tangent vectors in $\mathrm{T}_U\mathrm{St}(m, m-n)$ have the orthogonal decomposition $\dot{U} = U\dot{X} \oplus U^\perp\dot{Y}$, where $\dot{X} \in \mathrm{Skew}(m-n)$ is a skew-symmetric matrix of order $m-n$; $U^\perp$ is a matrix whose columns contain an orthonormal basis of the span of $U$ (hence, it contains a basis for the range of $A$); and $\dot{Y} \in \mathbb{R}^{n\times m-n}$. We can also orthogonally decompose $\dot{A} = \dot{P}U^T \oplus \dot{Q}(U^\perp)^T \in \mathbb{R}^{n\times m}$, where $\dot{P} \in \mathbb{R}^{n\times m-n}$ and $\dot{Q} \in \mathbb{R}^{n\times n}$.

By exploiting these decompositions and plugging them into eq. (3), we obtain

$$0 = A\dot{U} + \dot{P} \quad \text{so that} \quad -(AU^\perp)^{-1}\dot{P} = \dot{Y}.$$

Herein, we used the assumption that the range of $A$ is exactly the span of $U^\perp$, so that $AU^\perp$ is an $n \times n$ invertible matrix. (As an operator, this is $A$ restricted to its range.) We also find that

$$\dot{U}U^T + U\dot{U}^T = U^\perp\dot{Y}U^T + U\dot{Y}^T(U^\perp)^T = \begin{bmatrix} U & U^\perp \end{bmatrix} \begin{bmatrix} 0 & \dot{Y}^T \\ \dot{Y} & 0 \end{bmatrix} \begin{bmatrix} U & U^\perp \end{bmatrix}^T.$$

The first equality holds because $\dot{X}$ is skew symmetric, so $U\dot{X}U^T + U\dot{X}^TU^T = 0$. Consequently, for any unitarily invariant norm, we find

$$\|\dot{U}U^T + U\dot{U}^T\| = \left\| \begin{bmatrix} 0 & \dot{Y}^T \\ \dot{Y} & 0 \end{bmatrix} \right\|.$$

Suppose now that we take the Finsler metric $\|\cdot\|_F$ on the input space $\mathcal{I} = \mathbb{R}^{n\times m}$ and the Finsler metric $\|\cdot\|_F/\sqrt{2}$ on the output manifold $\mathcal{O} = \mathrm{Gr}(m-n, V)$.

The condition number at a matrix $A \in \mathbb{R}^{n\times m}$ that has a kernel of the expected dimension $m-n$ is

$$\kappa[\ker](A) = \sup_{\dot{A}\in\mathbb{R}^{n\times m}} \frac{\frac{1}{\sqrt{2}}\|\dot{U}U^T + U\dot{U}^T\|_F}{\|\dot{A}\|_F} = \frac{1}{\sqrt{2}} \sup_{\substack{\dot{P}\in\mathbb{R}^{n\times m-n},\\ \dot{Q}\in\mathbb{R}^{n\times n}}} \frac{\sqrt{2}\|\dot{Y}\|_F}{\sqrt{\|\dot{P}\|_F^2 + \|\dot{Q}\|_F^2}}.$$

In the last step we exploited the orthogonal decomposition of $\dot{A}$. Note that $\dot{Q}$ only appears in the denominator, since the numerator equals $\|(AU^\perp)^{-1}\dot{P}\|_F$, so that the optimal value of $\dot{Q}$ is obtained when $\|\dot{Q}\|_F = 0$. Then,

$$\kappa[\ker](A) = \sup_{\dot{P}\in\mathbb{R}^{n\times m-n}} \frac{\|(AU^\perp)^{-1}\dot{P}\|_F}{\|\dot{P}\|_F} = \|(AU^\perp)^{-1} \otimes I\|_2 = \sigma_n^{-1}(A).$$

The last step used that $\sigma_i(A) = \sigma_i(AU^\perp)$ for $i = 1, \ldots, n$, as $AU^\perp$ is simply the restriction of $A$ to its range. The second equality used the well known fact that $\mathrm{vec}(ABC^T) = (A \otimes C)\mathrm{vec}(B)$, where the vectorization uses the lexicographic order and $\otimes$ denotes the Kronecker product of matrices; see [10].[4]

Finally, if we choose the natural Finsler metric $\|\cdot\|_F/\|A\|_2$ to measure relative errors in the input space $\mathcal{I}$ at $A$, then the corresponding relative condition number

$$\mu[\ker](A) = \frac{\sigma_1(A)}{\sigma_n(A)} = \|A\|_2\|A^\dagger\|_2$$

is the (sometimes called "effective") matrix condition number $\kappa(A)$. The notation $A^\dagger$ denotes the so-called Moore–Penrose pseudoinverse of $A$ in numerical linear algebra (or simply the inverse of $A$ restricted to $U^\perp$). Note that errors are automatically measured in a relative sense on the output space because the Grassmannian is compact.

[4]In fact, the equation may be taken as the definition of vec.

## 4. Condition in bilinear algebra: Matrix multiplication

The next computational problem I consider is the multiplication of two real nonzero matrices. It is the bilinear map

$$\cdot : \mathbb{R}_*^{m\times n} \times \mathbb{R}_*^{n\times p} \to \mathbb{R}^{m\times p}, \quad (A,B) \mapsto AB.$$

It can be seen that its derivative is

$$\mathrm{d}_{(A,B)}\cdot : \mathbb{R}^{m\times n} \times \mathbb{R}^{n\times p}, \quad (\dot{A},\dot{B}) \mapsto \dot{A}B + A\dot{B}.$$

If we take the standard basis matrices $e_i e_j^T$ in lexicographic order, the matrix of this derivative can be expressed, as is well known [8, Chapter 4], as

$$J_{\alpha,\beta} := \mathrm{d}_{(A,B)}\cdot = \begin{bmatrix} \alpha I_{mn} \otimes B^T & \beta A \otimes I_{np} \end{bmatrix}.$$

The reason for introducing $\alpha$ and $\beta$ will become clear shortly.

The singular values of $J_{\alpha,\beta}$ correspond to the positive roots of the eigenvalues of

$$J_{\alpha,\beta} J_{\alpha,\beta}^T = \alpha^2 I_{mn} \otimes (B^T B) + \beta^2 (AA^T) \otimes I_{np}.$$

This particular sum is similar to the Kronecker sum in the literature, and its eigenvalues are easily shown to equal a particular linear combination of the eigenvalues of $B^T B$ and $AA^T$. The next lemma generalizes [8, Theorem 4.4.5] and can be proved along the same lines.

**Lemma 1.** *Let $M_k$ be square, diagonalizable matrices with eigenvalues $\lambda_{i_k}(M_k)$ and corresponding right eigenvector $\mathbf{v}_{i_k}(M_k)$. Consider the sum*

$$L = \sum_{i=1}^{m} \bigotimes_{k=1}^{d} \sum_{j=0}^{p} a_{ji} M_k^j.$$

*Then the eigenvalues of $L$ are*

$$\lambda_{i_1,i_2,\ldots,i_d}(L) = \sum_{i=1}^{m} \prod_{k=1}^{d} \sum_{j=0}^{p} a_{ji} \lambda_{i_k}^j(M_k)$$

*with corresponding eigenvector $\mathbf{v}_{i_1,\ldots,i_d} := \mathbf{v}_{i_1}(M_1) \otimes \cdots \otimes \mathbf{v}_{i_d}(M_d)$.*

*Proof.* Apply $L$ to $\mathbf{v}_{i_1,\ldots,i_d}$ and observe that the outcome is $\lambda_{i_1,\ldots,i_d}\mathbf{v}_{i_1,\ldots,i_d}$. □

Applying the lemma then yields

$$\lambda_{ij}(J_{\alpha,\beta} J_{\alpha,\beta}^T) = \alpha^2 \lambda_i(B^T B) + \beta^2 \lambda_j(AA^T).$$

In particular, by taking the largest eigenvalues of the positive semidefinite matrices $B^T B$ and $AA^T$, we find

$$\sigma_1(J_{\alpha,\beta}) = \sqrt{\sigma_1(\alpha B)^2 + \sigma_1(\beta A)^2}.$$

If we then take the product Frobenius norm as Finsler metric on $\mathcal{I} = \mathbb{R}^{m\times n} \times \mathbb{R}^{n\times p}$ and the Frobenius norm as Finsler metric on the output space $\mathcal{O} = \mathbb{R}^{m\times p}$, then we have the following isometries:

$$\left\| \begin{bmatrix} \dot{A}^T \\ \dot{B} \end{bmatrix} \right\|_F = \left\| \begin{bmatrix} \mathrm{vec}(\dot{A}) \\ \mathrm{vec}(\dot{B}) \end{bmatrix} \right\|_2 \quad \text{and} \quad \|(\mathrm{d}_{(A,B)}\cdot)(\dot{A},\dot{B})\|_F = \left\| J_{1,1} \begin{bmatrix} \mathrm{vec}(\dot{A}) \\ \mathrm{vec}(\dot{B}) \end{bmatrix} \right\|_2.$$

Hence, the absolute condition number of matrix multiplication is

$$\kappa[\cdot](A,B) = \sup_{\dot{\mathbf{x}} \in \mathbb{R}^{n(m+p)}} \frac{\|J_{1,1}\dot{\mathbf{x}}\|_2}{\|\dot{\mathbf{x}}\|_2} = \|J_{1,1}\|_2 = \sqrt{\|A\|_2^2 + \|B\|_2^2}.$$

A natural relative condition number is obtained by taking as Finsler metric on the input space, the product of Finsler metrics measuring relative errors in $\mathbb{R}^{m\times n}$ and $\mathbb{R}^{n\times p}$. Specifically, I consider

$$\|(\dot{A},\dot{B})\|_{\mathcal{I},(A,B)} = \sqrt{\frac{\|\dot{A}\|_F^2}{\|A\|_2^2} + \frac{\|\dot{B}\|_F^2}{\|B\|_2^2}} = \frac{\sqrt{\|B\|_2^2\|\dot{A}\|_F^2 + \|A\|_2^2\|\dot{B}\|_F^2}}{\|A\|_2\|B\|_2}.$$

In terms of the vectorizations of $\dot{A}$ and $\dot{B}$ we equivalently have

$$\|(\dot{A},\dot{B})\|_{\mathcal{I},(A,B)} = \frac{1}{\|A\|_2\|B\|_2}\left\|\begin{bmatrix}\|B\|_2^2\mathrm{vec}(\dot{A})\\ \|A\|_2^2\mathrm{vec}(\dot{B})\end{bmatrix}\right\|_2.$$

On the output space we can take the choice $\|\dot{X}\|_{\mathcal{O},X} = \frac{\|\dot{X}\|_F}{\sqrt{2}\|X\|_2}$. Then, the corresponding relative condition number is

$$\begin{aligned}\mu[\cdot](A,B) &= \sup_{(\dot{A},\dot{B})\in \mathrm{T}_{(A,B)}(\mathbb{R}^{m\times n}\times\mathbb{R}^{n\times p})} \frac{\|\dot{A}B + A\dot{B}\|_F/(\sqrt{2}\|AB\|_2)}{\sqrt{\|B\|_2^2\|\dot{A}\|_F^2 + \|A\|_2^2\|\dot{B}\|_F^2}/(\|A\|_2\|B\|_2)}\\
&= \frac{\|A\|_2\|B\|_2}{\sqrt{2}\|AB\|_2}\sup_{\dot{\mathbf{x}}\in\mathbb{R}^{n(m+p)}}\frac{\|J_{1,1}\dot{\mathbf{x}}\|_2}{\|\mathrm{diag}(\|B\|_2^2 I_{mn}, \|A\|_2^2 I_{np})\dot{\mathbf{x}}\|_2}\\
&= \frac{\|A\|_2\|B\|_2}{\sqrt{2}\|AB\|_2}\sup_{\dot{\mathbf{x}}\in\mathbb{R}^{n(m+p)}}\frac{\|J_{1,1}\mathrm{diag}(\|B\|_2^{-2} I_{mn}, \|A\|_2^{-2} I_{np})\dot{\mathbf{x}}\|_2}{\|\dot{\mathbf{x}}\|_2}\\
&= \frac{\|A\|_2\|B\|_2}{\sqrt{2}\|AB\|_2}\|J_{\|B\|_2^{-2},\|A\|_2^{-2}}\|_2.\end{aligned}$$

Since the spectral norm of $J_{\|B\|_2^{-2},\|A\|_2^{-2}}$ is $\sqrt{2}$, we obtain the ultimate expression

$$\mu[\cdot](A,B) = \frac{\|A\|_2\|B\|_2}{\|AB\|_2}.$$

This proves theorem 2.

**Remark 1.** It is a classic result that both standard and fast matrix multiplication algorithms satisfy a forward error bound of the form

$$\|\widetilde{C} - AB\| \le q(s)\|A\|\|B\|\cdot u,$$

where $\|\cdot\|$ is a unitarily invariant matrix norm, $\widetilde{C}$ is the numerically computed result of multiplying $A$ and $B$, $q$ is a polynomial (usually of low degree) in the problem size $s = \max\{m,n,p\}$, and $u$ is the unit roundoff [4]. When selecting the matrix spectral norm as norm, and dividing both sides by $\|AB\|_2$, if nonzero, we obtain

$$\frac{1}{\sqrt{\min\{m,p\}}}\frac{\|\widetilde{C}-AB\|_F}{\|AB\|_2} \le \frac{\|\widetilde{C}-AB\|_2}{\|AB\|_2} \le q(s)\frac{\|A\|_2\|B\|_2}{\|AB\|_2}\cdot u = \mu[\cdot](A,B)\cdot q(s)u.$$

This means that standard and fast matrix multiplication algorithms satisfying this bound are forward stable in the sense of [1, Definition 6] (with respect to the relative error Finsler metrics that I used in the definition of the relative condition number of matrix multiplication $\mu[\cdot]$) because the forward error is bounded by a polynomial in the input and output size multiplied by the condition number and the unit roundoff.

We also observe that that in the case of absolute errors, the fraction

$$\frac{\|A\|_2\|B\|_2}{\kappa[\cdot](A,B)} = \frac{\|A\|_2\|B\|_2}{\sqrt{\|A\|_2^2+\|B\|_2^2}}$$

is not bounded from above. In particular, fixing $A$ and letting $\|B\|_2 \to \infty$, we see that it tends to $\|A\|_2$. Hence, we only have the absolute forward error bound

$$\|\widetilde{C} - AB\|_F \le \|A\|_2 \kappa[\cdot](A,B) q(s) u$$

in this case, which is too large by a factor of $\|A\|_2$ to be called forward stable in the sense of [1, Definition 6].

## 5. Condition in multilinear algebra: Multilinear maps

A general multilinear map is of the form

$$\mathcal{A} : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d} \to \mathbb{R}^m, \ (\mathbf{x}_1, \ldots, \mathbf{x}_d) \mapsto \mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d),$$

where $\mathcal{A}$ is linear in each of its arguments.

The condition number of applying $\mathcal{A}$ can be computed as follows. By taking derivatives, we find that $(\dot{\mathbf{x}}_1, \ldots, \dot{\mathbf{x}}_d)$ is mapped to

$$\mathcal{A}(\dot{\mathbf{x}}_1, \mathbf{x}_2, \ldots, \mathbf{x}_d) + \cdots + \mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_{d-1}, \dot{\mathbf{x}}_d) = \begin{bmatrix} A_1 & \cdots & A_d \end{bmatrix} \begin{bmatrix} \dot{\mathbf{x}}_1 \\ \vdots \\ \dot{\mathbf{x}}_d \end{bmatrix},$$

where $A_i = [\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, e_j, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_d)]_{j=1}^{n_i} \in \mathbb{R}^{m \times n_i}$, $i = 1, \ldots, d$, denote the partial contractions. Let us choose the product Euclidean norm on the input space $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d}$, which equals the Euclidean norm on $\mathbb{R}^{n_1 + \cdots + n_d}$, and the Euclidean norm on the output space. It then follows from theorem 1 that the corresponding absolute condition number is

$$\kappa[\mathcal{A}](\mathbf{x}_1, \ldots, \mathbf{x}_d) = \left\| \begin{bmatrix} A_1 & \cdots & A_d \end{bmatrix} \right\|_2.$$

Let us choose the Finsler metrics $\|\cdot\|/\|\mathbf{x}_i\|$ on $\mathrm{T}_{\mathbf{x}_i}\mathbb{R}^{n_i}$ ($\mathbf{x}_i \neq 0$ of course) and then we take their product metric to measure errors on the input space. Equip the output space with the Finsler metric $\|\cdot\|/(\sqrt{d}\|\mathbf{y}\|)$ on $\mathrm{T}_{\mathbf{y}}\mathbb{R}^m$. Then, by reparameterizing, we have

$$\frac{1}{\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|} \sup_{\substack{\dot{\mathbf{x}}_i \in \mathbb{R}^{n_i}, \\ i=1,\ldots,d}} \frac{\left\| \begin{bmatrix} A_1 & \cdots & A_d \end{bmatrix} \begin{bmatrix} \dot{\mathbf{x}}_1 \\ \vdots \\ \dot{\mathbf{x}}_d \end{bmatrix} \right\|}{\sqrt{d} \left\| \begin{bmatrix} \dot{\mathbf{x}}_1/\|\mathbf{x}_1\| \\ \vdots \\ \dot{\mathbf{x}}_d/\|\mathbf{x}_d\| \end{bmatrix} \right\|} = \frac{\| \begin{bmatrix} \|\mathbf{x}_1\| A_1 & \cdots & \|\mathbf{x}_d\| A_d \end{bmatrix} \|_2}{\sqrt{d} \|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|}.$$

Consequently, we obtain the following relative condition number of applying the multilinear map $\mathcal{A}$ with respect to these Finsler metrics:

$$\mu[\mathcal{A}](\mathbf{x}_1, \ldots, \mathbf{x}_d) = \frac{\| \begin{bmatrix} \|\mathbf{x}_1\| A_1 & \cdots & \|\mathbf{x}_d\| A_d \end{bmatrix} \|_2}{\sqrt{d} \|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|}.$$

It is interesting to note what happens when $d = 1$, reducing to the linear case. In this case, the formula for the absolute condition number reduces to $\kappa[\mathcal{A}](\mathbf{x}_1) = \|\mathcal{A}\|_2$, coinciding with the absolute condition number of linear maps (when the same norms as above were chosen). And the relative condition number becomes $\mu[\mathcal{A}](\mathbf{x}_1) = \frac{\|\mathcal{A}\|_2 \|\mathbf{x}_1\|_2}{\|\mathcal{A}\mathbf{x}_1\|_2}$, which is precisely the same as in the case of linear maps (when the same Finsler metrics as above are chosen).

One may also be tempted to compare the case $d = 1$ of multilinear maps with the condition numbers obtained for matrix multiplication. At first sight something unusual may be noted. While the relative condition numbers coincide, $\mu[\mathcal{A}](\mathbf{x}_1) = \mu[\cdot](\mathcal{A}, \mathbf{x}_1)$, the absolute condition numbers do not:

$$\kappa[\mathcal{A}](\mathbf{x}_1) = \|\mathcal{A}\|_2 \neq \sqrt{\|\mathcal{A}\|_2 + \|\mathbf{x}_1\|_2} = \kappa[\cdot](\mathcal{A}, \mathbf{x}_1).$$

This is not a contraction, however. The reason is that the matrix condition number allows perturbations in both $\mathcal{A}$ and $\mathbf{x}_1$, while the condition number of applying the (multi)linear map $\mathcal{A}$ only allows perturbations in the argument $\mathbf{x}_1$. The correct way to apply the above result for multilinear maps is to take $d = 2$ and consider the multilinear map

$$M : \mathbb{R}^{mn} \times \mathbb{R}^{np} \to \mathbb{R}^{mp}, \ (\operatorname{vec}(A), \operatorname{vec}(B)) \mapsto \operatorname{vec}(AB)$$

Taking metrics as above, this approach, it can be verified as an exercise, results in

$$\kappa[\cdot](A, B) = \kappa[M](\operatorname{vec}(A), \operatorname{vec}(B)),$$
$$\mu[\cdot](A, B) = \mu[M](\operatorname{vec}(A), \operatorname{vec}(B)).$$

*Bounds.* Recall that the spectral norm of a tensor is defined as

$$\|\mathcal{A}\|_2 := \sup_{\substack{\mathbf{u}_i \in \mathbb{S}^{n_i - 1}, \\ i=1,\ldots,d}} \|\mathcal{A}(\mathbf{u}_1, \ldots, \mathbf{u}_d)\|.$$

Then it follows from the multilinearity of $\mathcal{A}$ that

$$\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\| \le \|\mathcal{A}\|_2 \|\mathbf{x}_1\| \cdots \|\mathbf{x}_d\|$$

for all $\mathbf{x}_i \in \mathbb{R}^{n_i}$. We will also use the standard fact

$$\left\| \begin{bmatrix} X_1 & \cdots & X_d \end{bmatrix} \right\|_2 = \sqrt{\lambda_1(X_1 X_1^T + \cdots + X_d X_d^T)} \le \sqrt{\|X_1\|_2^2 + \cdots + \|X_d\|_2^2}.$$

The absolute condition number can be bounded as follows.

$$\max_{1 \le i \le d} \|A_i\|_2 \le \kappa[\mathcal{A}](\mathbf{x}_1, \ldots, \mathbf{x}_d) \le \sqrt{\sum_{i=1}^d \|A_i\|_2^2} \le \sqrt{d} \max_{1 \le i \le d} \|A_i\|_2.$$

The relative condition number can be bounded from below by

$$\frac{1}{\sqrt{d}\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|} \max_{1 \le i \le d} \|\mathbf{x}_i\| \|A_i\|_2 \le \mu[\mathcal{A}](\mathbf{x}_1, \ldots, \mathbf{x}_d)$$

and from above by

$$\begin{aligned} \mu[\mathcal{A}](\mathbf{x}_1, \ldots, \mathbf{x}_d) &\le \frac{1}{\sqrt{d}\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|} \sqrt{\sum_{i=1}^d \|\mathbf{x}_i\|^2 \|A_i\|_2^2} \\ &\le \frac{1}{\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|} \max_{1 \le i \le d} \|\mathbf{x}_i\| \|A_i\|_2 \qquad \le \frac{\|\mathcal{A}\|_2 \|\mathbf{x}_1\| \cdots \|\mathbf{x}_d\|}{\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|}. \end{aligned}$$

This proves the following nice formula.

**Proposition 1.** *Let $\mathbf{x}_i \in \mathbb{R}^{n_i}$ be nonzero vectors for $i = 1, \ldots, d$. The relative condition number of applying a multilinear map $\mathcal{A}$ to these vectors satisfies:*

$$\mu[\mathcal{A}](\mathbf{x}_1, \ldots, \mathbf{x}_d) = \frac{\left\| \begin{bmatrix} \|\mathbf{x}_1\| A_1 & \cdots & \|\mathbf{x}_d\| A_d \end{bmatrix} \right\|_2}{\sqrt{d}\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|} \le \frac{\|\mathcal{A}\|_2 \|\mathbf{x}_1\| \cdots \|\mathbf{x}_d\|}{\|\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_d)\|},$$

*where $A_i = [\mathcal{A}(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, e_j, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_d)]_{j=1}^{n_i} \in \mathbb{R}^{m \times n_i}$, $i = 1, \ldots, d$ are the partial contractions and $\|\mathcal{A}\|_2$ is the tensor spectral norm.*

## References

1. C. Beltrán, V. Noferini, and N. Vannieuwenhoven, *When can forward stable algorithms be composed stably?*, arXiv:2109.10610 (2022).
2. L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*, Springer-Verlag, New York, 1998.
3. P. Bürgisser and F. Cucker, *Condition: The Geometry of Numerical Algorithms*, Springer, Heidelberg, 2013. MR 3098452
4. J. W. Demmel, I. Dumitriu, O. Holtz, and R. Kleinberg, *Fast matrix multiplication is stable*, Numerische Mathematik **106** (2007), no. 2, 199–224.
5. A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM Journal on Matrix Analysis and Applications **20** (1998), no. 2, 303–353.
6. W. Greub, *Multilinear Algebra*, 2 ed., Springer-Verlag, 1978.
7. N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2 ed., SIAM, Philadelphia, PA, 2002.
8. R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, 1991.
9. J. M. Lee, *Introduction to Smooth Manifolds*, 2 ed., Springer, New York, USA, 2013.
10. L.-H. Lim, *Tensors in computations*, Acta Numerica **30** (2021), 555–764.
11. P. Petersen, *Riemannian geometry*, second ed., Graduate Texts in Mathematics, vol. 171, Springer, New York, 2006.
12. J. D. Pryce, *A new measure of relative error for vectors*, SIAM Journal on Numerical Analysis **21** (1984), no. 1, 202–215.
13. J. R. Rice, *A theory of condition*, SIAM Journal on Numerical Analysis **3** (1966), no. 2, 287–310.