Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 08–43

USING DUAL TECHNIQUES TO DERIVE COMPONENTWISE AND MIXED CONDITION NUMBERS FOR A LINEAR FUNCTIONAL OF A LINEAR LEAST SQUARES SOLUTION

MARC BABOULIN AND SERGE GRATTON

Abstract: We prove duality results for adjoint operators and product norms in the framework of Euclidean spaces. We show how these results can be used to derive condition numbers especially when perturbations on data are measured componentwise relatively to the original data. We apply this technique to obtain formulas for componentwise and mixed condition numbers for a linear functional of a linear least squares solution. These expressions are closed when perturbations of the solution are measured using a componentwise norm or the infinity norm and we get an upper bound for the Euclidean norm.

Keywords: dual norm, adjoint operator, componentwise perturbations, condition number, linear least squares.
AMS Subject Classification (2000): 65F35.

1. Introduction.

The condition number of a problem measures the effect on the solution of small changes in the data and this quantity depends on the norms chosen for measuring perturbations on data and solution. A commonly used approach consists of measuring these perturbations in their globality using classical norms (e.g $\|\cdot\|_p,\ p = 1, 2, \infty$ or $\|\cdot\|_F$) resulting in so-called normwise condition numbers. But as mentioned in [20], using norms for measuring matrix perturbations has several drawbacks. First, it does not give information about how the perturbation is distributed among the data. Second, as pointed out in [7, p. 31], errors can be overestimated by a normwise sensitivity analysis when the problem is badly scaled.

In the *componentwise* analysis, perturbations are measured using metrics that take into account the structure of the matrix like sparsity or scaling. With such metrics, we could expect to minimize the amplification of perturbations resulting in a minimal condition number. Componentwise metrics are well suited for that because the perturbations on data are measured relatively to a given tolerance for each component of the data. For instance, if E

Received September 10, 2008.

is the matrix whose nonnegative entries are the individual tolerance for errors in components of A, then the componentwise relative error bound ω for a perturbation ΔA will be such that $|\Delta A| \leq \omega E$ (here partial ordering and absolute values must be understood component by component). A common choice for E is $E = |A|$.

Componentwise error analysis that provide us with exact expressions or bounds for componentwise condition numbers can be found for example in [5, 8, 19, 25, 27, 28, 29] for linear systems and in [4, 6, 9, 19, 23] for linear least squares. In particular, componentwise backward errors are commonly used as stopping criteria in iterative refinement for solving linear systems (see e.g [3]) or linear least squares (see e.g [10]).

For the full rank linear least squares problem (LLSP), generalizing [16], [2] gives exact formulas for the conditioning of a *linear transformation* of the LLSP solution when the perturbations of data and solution are measured normwise. Our objective in this paper is to obtain similar quantities when perturbations on data are, contrary to [2], measured componentwise and the perturbations on the solution are measured either componentwise or normwise resulting in respectively componentwise and mixed condition numbers.

In [17], a technique is presented to compute or estimate condition numbers using adjoint formulas. The results are presented in Banach spaces, and make use of the corresponding duality results in order to derive normwise condition numbers. In our paper, we show that these dual techniques are easy and helpful when presented in the framework of Euclidean spaces. In particular, as also mentioned in [17], they enable us to derive condition numbers by maximizing a linear functional over a space of smaller dimension than the data space.

We show in this paper that dual techniques can be used to derive condition numbers when perturbations on the data are measured componentwise and we apply this method to LLSP. We propose exact formulas for the conditioning of $L^T x$, linear functional of the LLSP solution when perturbations on data are measured componentwise and perturbations on the solution are measured either componentwise or normwise. Studying the condition number of $L^T x$ is relevant for instance when there is a dispararity between the conditioning of the solution components or when the computation of the least squares solution involves auxiliary variables without physical significance.
The situations of common interest correspond to the cases where L is the identity matrix (condition number of an LLSP solution), a canonical vector

(condition number of one solution component), or a projector, when we are interested in the sensitivity of a subset of the solution components. The conditioning of a nonlinear function of an LLSP solution can also be obtained by replacing in the condition number expression L^T by the Jacobian matrix at the solution. When L is the identity matrix and when perturbations on the solution are measured using the infinity norm or a componentwise norm, we obtain the exact formula given in [9]. By considering the special case where we have a residual equal to zero, we obtain componentwise and mixed condition numbers for $L^T x$ where x is the solution of a consistent linear system. When L is the identity matrix, these quantities recover the expressions known from [21, p. 123] and [18].

2. Deriving condition numbers using dual techniques.

2.1. Preliminary results on dual norms and adjoint operators. We consider a linear mapping $J : E \to G$ where the Euclidean spaces E and G are equipped respectively with any norms $\|.\|_E$, $\|.\|_G$ and scalar products $< .,. >_E$ and $< .,. >_G$. Note that the norms $\|.\|_E$ and $\|.\|_G$ may not, and will not in general be, the particular norms induced by the scalar products $< .,. >_E$ and $< .,. >_G$.

Definition 1. *The adjoint operator of J, $J^* : G \to E$ is defined by*

$$< y, Jx >_G = < J^* y, x >_E,$$

where $(x, y) \in E \times G$.
We also define the dual norm $\|.\|_{E*}$ of $\|.\|_E$ by

$$\|x\|_{E*} = \max_{u \neq 0} \frac{< x, u >_E}{\|u\|_E},$$

and define similarly the dual norm $\|.\|_{G*}$.

For the common vector norms, the dual norms with respect to the canonical scalar product in $\mathbb{R}^n$ are well-known, and are given by:

$$\|\cdot\|_{1*} = \|\cdot\|_\infty \text{ and } \|\cdot\|_{\infty*} = \|\cdot\|_1 \text{ and } \|\cdot\|_{2*} = \|\cdot\|_2.$$

For the matrix norms in $\mathbb{R}^{m\times n}$ with respect to the scalar product $< A, B >=$ trace$(A^T B)$, we have $\|A\|_{2*} = \|\sigma(A)\|_1$ (Lemma 3.5 of [30, p. 78]), and since trace$(A^T A) = \|A\|_F^2$ we also have that $\|A\|_{F*} = \|A\|_F$.

For the linear applications mapping E to G, we denote by $\|.\|_{E,G}$ the operator norm induced by the norms $\|.\|_E$ and $\|.\|_G$. Likewise, the norm $\|.\|_{G*,E*}$ is the operator norm for linear applications mapping G to E and induced by the dual norms $\|.\|_{G*}$ and $\|.\|_{E*}$. Then we have the following theorem.

Theorem 1.

$$\|J\|_{E,G} = \|J^*\|_{G*,E*}$$

Proof:

$$\begin{aligned}
\|J\|_{E,G} &= \max_{x\in E} \frac{\|Jx\|_G}{\|x\|_E} \\
&= \max_{x\in E, u\in G} \frac{< Jx, u >_G}{\|u\|_{G*}\|x\|_E} \quad \text{we use the “duality theorem” [22, p. 287]} \\
&= \max_{u\in G} \frac{1}{\|u\|_{G*}} \max_{x\in E} \frac{< x, J^*u >_E}{\|x\|_E} \\
&= \max_{u\in G} \frac{\|J^*u\|_{E*}}{\|u\|_{G*}} \\
&= \|J^*\|_{G*,E*}.
\end{aligned}$$

■

As mentioned in [17], it may be interesting to compute $\|J^*\|_{G*,E*}$ (instead of $\|J\|_{E,G}$) when $G*$ is an Euclidean space of lower dimension than E because it implies a maximization over a space of smaller dimension.
We now consider a product space $E = E_1 \times \cdots \times E_p$ where each Euclidean space E_i is equipped with the norm $\|.\|_{E_i}$ and the scalar product $< .,. >_{E_i}$.
In E, we consider the following scalar product

$$< (u_1, \ldots, u_p), (v_1, \ldots, v_p) >=< u_1, v_1 >_{E_1} + \cdots + < u_p, v_p >_{E_p},$$

and the product norm

$$\|(u_1, \ldots, u_p)\|_\nu = \nu(\|u_1\|_{E_1}, \ldots, \|u_p\|_{E_p}),$$

where ν is an absolute norm on $\mathbb{R}^p$ (i.e such that $\nu(|x|) = \nu(x)\ \forall x \in \mathbb{R}^p$, [24, p. 367]). We denote by ν_* the dual of ν with respect to the canonical inner-product of $\mathbb{R}^p$ and we are interested in determining the dual $\|.\|_{\nu*}$ of the product norm $\|.\|_\nu$ with respect to the scalar product of E. Then we have the following result.

Theorem 2. *The dual of the product norm can be expressed as*

$$\|(u_1,\dots,u_p)\|_{\nu*} = \nu_*(\|u_1\|_{E_1*},\dots,\|u_p\|_{E_p*}).$$

Proof: From $\|u_i\|_{E_i*} = \max_{v_i\neq 0} \frac{<u_i,v_i>_{E_i}}{\|v_i\|_{E_i}}$, we have $\forall v_i \in E_i \quad < u_i, v_i >_{E_i} \leq \|u_i\|_{E_i*}\|v_i\|_{E_i}$ and then

$$\begin{aligned}
\|(u_1,\dots,u_p)\|_{\nu*} &= \max_{\|(v_1,\dots,v_p)\|_\nu=1} \sum_{i=1}^{p} < u_i, v_i >_{E_i} \\
&\leq \max_{\|(v_1,\dots,v_p)\|_\nu=1} \sum_{i=1}^{p} \|u_i\|_{E_i*}\|v_i\|_{E_i} \\
&= \max_{\nu(\|v_1\|_{E_1},\dots,\|v_p\|_{E_p})=1} \begin{pmatrix} \|u_1\|_{E_1*} \\ \vdots \\ \|u_p\|_{E_p*} \end{pmatrix}^T \begin{pmatrix} \|v_1\|_{E_1} \\ \vdots \\ \|v_p\|_{E_p} \end{pmatrix} \\
&= \nu_*(\|u_1\|_{E_1*},\dots,\|u_p\|_{E_p*}).
\end{aligned}$$

So, we have shown that $\nu_*(\|u_1\|_{E_1*},\dots,\|u_p\|_{E_p*})$ is an upper-bound for the dual of the product norm.
Now let $w_1,\dots,w_p$ be nonzero vectors such that $\forall i, < u_i, w_i >_{E_i} = \|u_i\|_{E_i*}\|w_i\|_{E_i}$ (i.e choose w_i that attains the maximum in the definition of the dual norm $\|u_i\|_{E_i*}$). Then

$$\nu_*(\|u_1\|_{E_1*},\dots,\|u_p\|_{E_p*}) = \max_{\nu(\alpha_1\|w_1\|_{E_1},\dots,\alpha_p\|w_p\|_{E_p})=1} \begin{pmatrix} \|u_1\|_{E_1*} \\ \vdots \\ \|u_p\|_{E_p*} \end{pmatrix}^T \begin{pmatrix} \alpha_1\|w_1\|_{E_1} \\ \vdots \\ \alpha_p\|w_p\|_{E_p} \end{pmatrix}$$

is attained for a particular $(\alpha'_1,\dots,\alpha'_p)$ such that

$$\nu_*(\|u_1\|_{E_1*},\dots,\|u_p\|_{E_p*}) = \sum_{i=1}^{p} \alpha'_i < u_i, w_i >_{E_i},$$

with $\nu(\alpha'_1\|w_1\|_{E_1},\dots,\alpha'_p\|w_p\|_{E_p}) = 1$. Using the fact that ν is an absolute norm, we get

$$\begin{aligned}
\|(\alpha'_1 w_1,\dots,\alpha'_p w_p)\|_\nu &= \nu(\|\alpha'_1 w_1\|_{E_1},\dots,\|\alpha'_p w_p\|_{E_p}) \\
&= \nu(|\alpha'_1|\|w_1\|_{E_1},\dots,|\alpha'_p|\|w_p\|_{E_p}) \\
&= \nu(\alpha'_1\|w_1\|_{E_1},\dots,\alpha'_p\|w_p\|_{E_p}) \\
&= 1.
\end{aligned}$$

Thus

$$\begin{aligned}\|(u_1,\ldots,u_p)\|_{\nu*} &= \max_{\|(v_1,\ldots,v_p)\|_{\nu}=1} \sum_{i=1}^{p} < u_i, v_i >_{E_i} \\ &\geq \sum_{i=1}^{p} < u_i, \alpha_i' w_i >_{E_i} \\ &= \nu_*(\|u_1\|_{E_1*},\ldots,\|u_p\|_{E_p*}),\end{aligned}$$

which concludes the proof. ■

2.2. Application to condition numbers. We represent here a given problem as a mapping g defined by $x = g(y)$ where $x \in G$ is the solution of the problem corresponding to the data $y \in E$. The data space E and the solution space G are Euclidean spaces equipped respectively with the norms $\|\cdot\|_E$ and $\|\cdot\|_G$. Then the condition number of the problem is a measure of the sensitivity of the mapping g to perturbations.

Following [26], if g is Fréchet-differentiable in a neighborhood of y, the **absolute condition number** of g at the point $y \in E$ is the quantity $K(y)$ defined by

$$K(y) = \|g'(y)\|_{E,G},$$

where $\|\cdot\|_{E,G}$ is the operator norm induced by the norms $\|\cdot\|_E$ and $\|\cdot\|_G$. Then we also have

$$K(y) = \max_{\|z\|_E=1} \|g'(y).z\|_G. \quad (1)$$

If $g(y)$ is nonzero, we can define the **relative condition number** of g at $y \in E$ as

$$K^{(rel)}(y) = K(y)\|y\|_E/\|g(y)\|_G. \quad (2)$$

The expression of $K(y)$ corresponds to the operator norm of the linear operator $g'(y)$. Then using Theorem 1 and with the same notations as Section 2.1, $K(y)$ can be expressed as

$$K(y) = \max_{\|\Delta y\|_E=1} \|g'(y).\Delta y\|_G = \max_{\|x\|_{G*}=1} \|g'(y)^*.x\|_{E*}. \quad (3)$$

We can summarize the method for deriving condition numbers using dual techniques as follows:

(1) determine the derivative $g'(y)$ of the mapping that represents our problem,

(2) choose the norms $\|\cdot\|_E$ and $\|\cdot\|_G$ respectively on the solution and the data spaces and determine their dual norms,

(3) determine the adjoint operator $g'(y)^*$ of the linear operator $g'(y)$,

(4) compute $K(y) = \max_{\|x\|_{G*}=1} \|g'(y)^*.x\|_{E*}$.

Let us now consider the case where we use a componentwise metric on a data space $E = \mathbb{R}^n$. For a given $y \in \mathbb{R}^n$, we denote by E_Y the subset of all elements $\Delta y \in \mathbb{R}^n$ with $\Delta y_i = 0$ whenever $y_i = 0$. Then in a componentwise perturbation analysis, we measure the perturbation $\Delta y \in E_Y$ of y using the norm

$$\|\Delta y\|_c = \min\{\omega, |\Delta y_i| \leq \omega |y_i|, i = 1, \ldots, n\}. \tag{4}$$

$\|\cdot\|_c$ is called the **componentwise relative norm** with respect to y.
As mentioned in [12], we can extend the definition in Equation (4) to the case where $\Delta y_i \neq 0$ while $y_i = 0$ by having the convention $\|\Delta y\|_c = \infty$ for those Δy.

Let us determine the dual norm of the componentwise norm. First we observe that, for any $\Delta y \in E_Y$, we have

$$\|\Delta y\|_c = \max\left\{\frac{|\Delta y_i|}{|y_i|},\ y_i \neq 0\right\} = \left\|\left(\frac{|\Delta y_1|}{|y_1|},\ y_i \neq 0\right)\right\|_\infty.$$

Then we can apply Theorem 2 by considering the product space $E = \mathbb{R}^n$, with each $E_i = \mathbb{R}$, $\nu = \|\cdot\|_\infty$ and $\|\Delta y_i\|_{E_i} = \frac{|\Delta y_i|}{|y_i|}$, if $y_i \neq 0$. We have

$$\|\Delta y_i\|_{E_i*} = \max_{z \neq 0} \frac{|\Delta y_i.z|}{\|z\|_{E_i}} = \max_{z\neq 0} |\Delta y_i.z|/(|z|/|y_i|) = |\Delta y_i||y_i|,$$

and also $\|\cdot\|_{\infty*} = \|\cdot\|_1$. Then we get

$$\|\Delta y\|_{c*} = \|(|\Delta y_1||y_1|, \ldots, |\Delta y_n||y_n|)\|_1. \tag{5}$$

If there are zero components in y, we observe that, due to the condition $\|\Delta y\|_E = 1$ in Equation (1), the definition of $K(y)$ is the same whether Δy is in E_Y or not. Indeed, if $\Delta y \notin E_Y$, then with the convention given previously we have $\|\Delta y\|_c = \infty$ and the perturbation Δy is not taken into account in the computation of $K(y)$. As a result, the zero components of y should not be explicitely excluded as data.
Using Equation (3), $K(y)$ will be obtained with

$$K(y) = \max_{\|x\|_{G*}=1} \|g'(y)^*.x\|_{c*}. \tag{6}$$

where $\|\cdot\|_{c*}$ is expressed by Equation (5).
Note that the norm on the solution space G has not been chosen yet. Following the terminology given in [13] and also used in [9], $K(y)$ is referred to as componentwise (resp. mixed) condition number when $\|\cdot\|_G$ is componentwise (resp. normwise). In Section 3, we consider different norms for the solution space but the norm on the data space is always componentwise.

3. Componentwise and mixed condition numbers for a linear functional of an LLSP solution.

3.1. Least squares conditioning. We consider the linear least squares problem $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2$, where $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$ is a matrix of full column rank n. Then the unique solution x is expressed by $x = (A^TA)^{-1}A^Tb = A^\dagger b$, where $A^\dagger$ denotes the pseudo-inverse of A. In the remainder, the matrix I is the identity matrix and e_i may denote the ith canonical vector of $\mathbb{R}^m$ or $\mathbb{R}^n$.
We study here the sensitivity of a linear functional of the LLSP solution to perturbations in the data A an b, which corresponds to the function

$$g(A, b) = L^T(A^TA)^{-1}A^Tb = L^TA^\dagger b,$$

where L is an $n \times k$ matrix, with $k \leq n$. In the most common case, L is the identity matrix (conditioning of the solution) but L can also be for instance a canonical vector of $\mathbb{R}^n$ if we are interested in the conditioning of one component of x. In the sequel, we suppose that L is not numerically perturbed. Since A has full rank n, g is continuously F-differentiable in a neighborhood of (A, b) and we denote by $J = g'(A, b)$ its derivative.
Let $B \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$. Using the chain rules of composition of derivatives, we get

$$g'(A, b).(B, c) =$$

$$L^T(A^TA)^{-1}B^T(b - A(A^TA)^{-1}A^Tb) - L^T(A^TA)^{-1}A^TB(A^TA)^{-1}A^Tb + L^TA^\dagger c$$

i.e

$$J(B, c) = g'(A, b).(B, c) = L^T(A^TA)^{-1}B^Tr - L^TA^\dagger Bx + L^TA^\dagger c,$$

where $r = b - Ax$ is the residual vector.

Remark 1. If we define $x(A,b) = A^{\dagger}b$, the case where $g(A,b) = h(x(A,b))$, with h being a differentiable nonlinear function mapping $\mathbb{R}^n$ to $\mathbb{R}^k$ is also covered because we have

$$g'(A,b).(B,c) = h'(A^{\dagger}b).(x'(A,b).(B,c)),$$

and L^T would correspond to the Jacobian matrix $h'(A^{\dagger}b)$.

We can find in [2] closed formulas, bounds and statistical estimates for the conditioning of $g(A,b)$ when perturbations on the data are measured normwise using the weighted norm

$$\left(\alpha^2 \|A\|^2_{\mathrm{F\ or\ 2}} + \beta^2 \|b\|_2^2\right)^{\frac{1}{2}}, \ \alpha, \beta > 0.$$

Here we are interested in the case where perturbations of A and b are measured componentwise.

3.2. Choice of norms. We consider the following norms and scalar products:

- for any vector u, $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are the vector norm corresponding to the classical definitions $\|\cdot\|_1 = \sum_i |u_i|$, $\|\cdot\|_2 = \left(\sum_i u_i^2\right)^{\frac{1}{2}}$ and $\|\cdot\|_\infty = \max_i |u_i|$.
- on the solution space $\mathbb{R}^k$, we use the scalar product $< x, y >= x^T y$, where the norm $\|\cdot\|$ can be $\|\cdot\|_2$, $\|\cdot\|_\infty$ or a componentwise norm with respect to the solution $L^T x$, and its dual norm is denoted by $\|\cdot\|_*$.
- on the data space $\mathbb{R}^{m\times n} \times \mathbb{R}^m$, we use the scalar product

$$< (A,b), (B,c) >= \mathrm{trace}(A^T B) + b^T c,$$

  and the componentwise relative norm (as given e.g in [21, p. 122]): $\|(\Delta A, \Delta b)\|_c = \min\{\omega, |\Delta A| \leq \omega|A|, |\Delta b| \leq \omega|b|\}$ where absolute values and inequalities between matrices or vectors are understood to hold componentwise.

3.3. Determination of the adjoint operator. The following proposition gives us the expression of the adjoint operator of J.

Proposition 1. *The adjoint of J, Fréchet derivative of a linear functional of the full rank least squares solution,*

$$\begin{array}{rcl} J \;:\; \mathbb{R}^{m\times n}\times\mathbb{R}^m & \longrightarrow & \mathbb{R}^k \\ (B,c) & \longmapsto & L^T(A^TA)^{-1}B^Tr - L^TA^\dagger Bx + L^TA^\dagger c \\ & & = J_1B + J_2c \end{array} \quad (7)$$

is

$$\begin{array}{rcl} J^* \;:\; \mathbb{R}^k & \longrightarrow & \mathbb{R}^{m\times n}\times\mathbb{R}^m \\ u & \longmapsto & (ru^TL^T(A^TA)^{-1} - A^{\dagger T}Lux^T, A^{\dagger T}Lu) \end{array} . \quad (8)$$

Proof: Using (7), we obtain for the first part of the adjoint of the derivative J,

$\forall u \in \mathbb{R}^k$, we have,

$$\begin{array}{rcl} < u, J_1B > & = & u^T(L^T(A^TA)^{-1}B^Tr - L^TA^\dagger Bx) \\ & = & \mathrm{trace}(L^T(A^TA)^{-1}B^Tru^T) - \mathrm{trace}(L^TA^\dagger Bxu^T) \\ & = & \mathrm{trace}(ru^TL^T(A^TA)^{-1}B^T) - \mathrm{trace}(xu^TL^TA^\dagger B) \\ & = & \mathrm{trace}((ru^TL^T(A^TA)^{-1})^TB) - \mathrm{trace}(xu^TL^TA^\dagger B) \\ & = & \mathrm{trace}(((ru^T(A^TA)^{-1})^T - xu^TL^TA^\dagger)B) \\ & = & < ru^TL^T(A^TA)^{-1} - A^{\dagger T}Lux^T, B > \\ & = & < {J_1}^*u, B > . \end{array}$$

For the second part of the adjoint of the derivative J, we have

$$\begin{array}{rcl} \forall u \in \mathbb{R}^k < u, J_2c > & = & u^TL^TA^\dagger c \\ & = & < A^{\dagger T}Lu, c > \\ & = & < {J_2}^*u, c > . \end{array}$$

■

As already mentioned in Section 2.1, the advantage of working with the adjoint J^* here is that the operator norm computation, involved in the condition number definition, implies a maximization over a vector space of dimension k, instead of a maximization over a vector space of dimension $mn+m$ for J. Indeed, using Equation (6), the condition number of L^Tx is given by

$$K(L,A,b) = \max_{\|(\Delta A,\Delta b)\|_c=1} \|J(\Delta A,\Delta b)\| = \max_{\|u\|_*=1} \|J^*(u)\|_{c*}.$$

3.4. Expression of the condition number. The following theorem provides us with an explicit formula for $K(L, A, b)$ thanks to the use of J^*. In the remainder, *vec* is the operator that stacks the columns of a matrix into a long vector, $\otimes$ denotes the Kronecker product of two matrices [15], and for any m-by-n matrix $Y = (y_{ij})$, D_Y denotes the diagonal matrix $diag(vec(Y)) = diag(y_{11}, \ldots, y_{m1}, \ldots, y_{1n}, \ldots, y_{mn})$.

Theorem 3. *The condition number of $L^T x$, linear functional of the full rank least squares solution, is expressed by*

$$K(L, A, b) = \| \, [V D_A, W D_b]^T \|_{*,1}, \text{ where}$$

$$V = (L^T (A^T A)^{-1}) \otimes r^T - x^T \otimes (L^T A^\dagger), \; W = L^T A^\dagger,$$

and $\|\cdot\|_{*,1}$ is the matrix norm subordinate to the vector norms $\|\cdot\|_*$ and $\|\cdot\|_1$ defined in Section 3.2.

Proof: If (Δa_{ij}) and (Δb_i) are the entries of ΔA and Δb then, using Equation (5), we have

$$\|(\Delta A, \Delta b)\|_{c*} = \sum_{i,j} |\Delta a_{ij}||a_{ij}| + \sum_i |\Delta b_i||b_i|.$$

Then, using Proposition 1, we get

$$\begin{aligned} \|J^*(u)\|_{c*} &= \sum_{j=1}^{n} \sum_{i=1}^{m} |a_{ij}||(r u^T L^T (A^T A)^{-1} - A^{\dagger T} L u x^T)_{ij}| + \sum_{i=1}^{m} |b_i||(A^{\dagger T} L u)_i| \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} |a_{ij}||(r_i {e_j}^T (A^T A)^{-1} - x_j {e_i}^T A^{\dagger T}) L u| + \sum_{i=1}^{m} |b_i||{e_i}^T A^{\dagger T} L u| \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} |v_{ij}^T a_{ij} u| + \sum_{i=1}^{m} |w_i^T b_i u|, \end{aligned}$$

where $v_{ij} = \left(r_i L^T (A^T A)^{-1} e_j - x_j L^T A^\dagger e_i \right) = \left(L^T (A^T A)^{-1} e_j r^T - x_j L^T A^\dagger \right) e_i$ and $w_i = L^T A^\dagger e_i$.

Note that, in the column vectors v_{ij} and w_i, the vectors e_j and e_i are canonical vectors from different spaces (respectively $\mathbb{R}^n$ and $\mathbb{R}^m$).

The quantities $v_{ij}^T a_{ij} u$ and $w_i^T b_i u$ are scalars and $\|J^*(u)\|_{c*}$ can be interpreted as the 1-norm of a vector as

$$\begin{aligned}
\|J^*(u)\|_{c*} &= \left\|(v_{11}a_{11},\ldots,v_{m1}a_{m1},\ldots,v_{1n}a_{1n},\ldots,v_{mn}a_{mn},w_1b_1,\ldots,w_mb_m)^T u\right\|_1 \\
&= \left\|[VD_A, WD_b]^T u\right\|_1,
\end{aligned}$$

where V is the k-by-mn matrix whose columns are the v_{ij} ordered in j first and W is the k-by-m matrix whose columns are the w_i.
V can be expressed as

$$\begin{aligned}
V &= \left(L^T(A^TA)^{-1}e_1r^T - x_1L^TA^\dagger,\ldots,L^T(A^TA)^{-1}e_nr^T - x_nL^TA^\dagger\right) \\
&= \left((L^T(A^TA)^{-1})\otimes r^T - x^T\otimes(L^TA^\dagger)\right),
\end{aligned}$$

and we also have $W = \left(L^TA^\dagger e_1,\ldots,L^TA^\dagger e_m\right) = L^TA^\dagger$.
Finally we get

$$\|J^*(u)\|_{c*} = \left\|[VD_A, WD_b]^T u\right\|_1,$$

and

$$K(L,A,b) = \max_{\|u\|_*=1}\left\| \ [VD_A, WD_b]^T u\right\|_1 = \| \ [VD_A, WD_b]^T\|_{*,1}.$$

■

Depending on the norm chosen for the solution space $\mathbb{R}^k$, we can have different expressions for $K(L,A,b)$. In the following section, we apply Theorem 3 to obtain expressions of $K(L,A,b)$ for some commonly used norms. Using the terminology given in Section 2.2, $K(L,A,b)$ will be referred to as mixed (resp. componentwise) condition number if the perturbations of the solution are measured normwise (resp. componentwise).

3.5. Condition number expressions for some norms on the solution space.

3.5.1. *Use of the infinity norm on the solution space.* If $\|\cdot\| = \|\cdot\|_\infty$, then $\|\cdot\|_* = \|\cdot\|_1$ and we have

$$K_\infty(L,A,b) = \left\|[VD_A, WD_b]^T\right\|_1 = \|[VD_A, WD_b]\|_\infty.$$

Then, with the notations used in the proof of Theorem 3, we get

$$\begin{aligned} K_\infty(L,A,b) &= \|(v_{11}a_{11},\dots,v_{m1}a_{m1},\dots,v_{1n}a_{1n},\dots,v_{mn}a_{mn},w_1b_1,\dots,w_mb_m)\|_\infty \\ &= \||v_{11}a_{11}|+\cdots+|v_{m1}a_{m1}|+\cdots+|v_{1n}a_{1n}|+\cdots+ \\ &\qquad +|v_{mn}a_{mn}|+|w_1b_1|+\cdots+|w_mb_m|\|_\infty \\ &= \||V|vec(|A|)+|W||b|\|_\infty, \end{aligned}$$

and thus

$$K_\infty(L,A,b) = \left\| \left|(L^T(A^TA)^{-1})\otimes r^T - x^T\otimes(L^TA^\dagger)\right| vec(|A|) + |L^TA^\dagger||b| \right\|_\infty . \tag{9}$$

Remark 2. For small problems, Matlab has a routine **kron** that enables us to compute $K_\infty(L,A,b)$ using the syntax:
Kinf=
norm(abs(kron(C,r')-kron(x',L'*Ap))*vec(abs(A))+abs(L'*pinv(A))*abs(b),inf),
with vec=inline('A(:)','A').
We also observe that $K_\infty(L,A,b)$ can also be written

$$K_\infty(L,A,b) = \left\| \sum_{j=1}^{n} [|v_{1j}|,\dots,|v_{mj}|]|A(:,j)| + |L^TA^\dagger||b| \right\|_\infty ,$$

and since $v_{ij} = L^T(A^TA)^{-1}(e_jr^T - x_jA^T)e_i$, we get

$$[|v_{1j}|,\dots,|v_{mj}|] = |L^T(A^TA)^{-1}[e_jr^T - x_jA^T]|.$$

Then we have

$$K_\infty(L,A,b) = \left\| \sum_{j=1}^{n} |L^T(A^TA)^{-1}(e_jr^T - x_jA^T)||A(:,j)| + |L^TA^\dagger||b| \right\|_\infty . \tag{10}$$

Equation (10) has the advantage to avoid forming the Kronecker products and then it requires less memory. Moreover, if the LLSP is solved using a direct method, the R factor of the QR decomposition of A (or equivalently in exact arithmetic, the Cholesky factor of A^TA) might be available and we have $(A^TA)^{-1} = R^{-1}R^{-T}$. Then the computation of $(A^TA)^{-1}(e_jr^T - x_jA^T)$ can be performed by solving successively 2 triangular systems with multiple right-hand sides and Equation (10) can be easily implemented using LAPACK [1] and Level 3 BLAS [11] routines.

When L is the identity matrix, the expression given in Equation (9) is the same as the one established in [9] (using the norm $\|\cdot\| = \frac{\|\cdot\|_\infty}{\|x\|_\infty}$ on the solution space). Note that bounds of $K_\infty(I, A, b)$ are also given in [7, p. 34] and in [21, p. 384].

When $m = n$ (case of a linear system $Ax = b$), we obtain, using the formulas $|A \otimes B| = |A| \otimes |B|$ and $vec(AYB) = (B^T \otimes A)vec(Y)$:

$$\begin{aligned} K_\infty(L, A, b) &= \left\| |x^T \otimes (L^T A^{-1})| vec(|A| + |L^T A^{-1}||b| \right\|_\infty \\ &= \left\| vec(|L^T A^{-1}||A||x|) + |L^T A^{-1}||b| \right\|_\infty . \end{aligned}$$

But $|L^T A^{-1}||A||x|$ is a vector and is then equal to its vec operator and then we obtain

$$K_\infty(L, A, b) = \left\| |L^T A^{-1}|(|A||x| + |b|) \right\|_\infty , \tag{11}$$

which generalizes the condition number given in [21, p. 123] to the case where L is not the identity (using the norm $\|\cdot\| = \frac{\|\cdot\|_\infty}{\|x\|_\infty}$ on the solution space).

3.5.2. *Use of the* 2*-norm on the solution space.* If $\|\cdot\| = \|\cdot\|_2$ then $\|\cdot\|_* = \|\cdot\|_2$ and, using Theorem 3, the mixed condition number of $L^T x$ is

$$K_2(L, A, b) = \|[VD_A, WD_b]^T\|_{2,1},$$

where $\|\cdot\|_{2,1}$ is the matrix norm subordinate to the vector norms $\|\cdot\|_2$ and $\|\cdot\|_1$.

As mentioned in [14, p. 56], we have for any matrix B,

$$\|B\|_{2,1} = \max_{\|u\|_2=1} \|Bu\|_1 = \|B\overline{u}\|_1$$

for some $\overline{u} \in \mathbb{R}^k$ having unit 2-norm. Since $\|\overline{u}\|_1 \leq \sqrt{k}\, \|\overline{u}\|_2$, we get

$$\|B\|_{2,1} = \|B\overline{u}\|_1 \leq \|\overline{u}\|_1 \|B\|_1 \leq \sqrt{k}\, \|B\|_1 .$$

Applying this inequality to $B = [VD_A, WD_b]^T$, we obtain,

$$K_2(L, A, b) \leq \sqrt{k} \left\| [VD_A, WD_b]^T \right\|_1 ,$$

and then we have the following upper bound for the mixed condition number of $L^T x$

$$K_2(L, A, b) \leq \sqrt{k}\; K_\infty(L, A, b), \tag{12}$$

and this upper bound can be computed using Equations (9) or (10).

3.5.3. *Use of a componentwise norm on the solution space.* We consider here a componentwise norm on the solution space defined by

$$\|u\| = \min \left\{\omega, |u_i| \leq \omega |(L^T x)_i|, i = 1, \ldots, k\right\}.$$

Then, following Equation (5), the dual norm is expressed by

$$\|u\|_* = \left\|(|(L^T x)_1||u_1|, \ldots, |(L^T x)_k||u_k|)\right\|_1.$$

With the convention mentioned in Section 2.2, we consider perturbations u in the solution space such that $u_i = 0$ whenever $(L^T x)_i = 0$. Let $D_{L^T x} = diag(\alpha_1, \ldots, \alpha_k)$ be the k-by-k diagonal matrix such that $\alpha_i = (L^T x)_i$ if $(L^T x)_i \neq 0$ and $\alpha_i = 1$ otherwise. Then if we apply Theorem 3 and if we perform the change of variable $u' = D_{L^T x} u$, the componentwise condition number of $L^T x$ is

$$\begin{aligned} K_c(L, A, b) &= \max_{\|u'\|_1 = 1} \left\|[V D_A, W D_b]^T D_{L^T x}^{-1} u'\right\|_1 \\ &= \left\|D_{L^T x}^{-1} [V D_A, W D_b]\right\|_\infty, \end{aligned}$$

that can be computed using the following variant of (10):

$$K_c(L, A, b) = \left\| \sum_{j=1}^{n} |D_{L^T x}^{-1} L^T (A^T A)^{-1} (e_j r^T - x_j A^T)||A(:, j)| + |D_{L^T x}^{-1} L^T A^\dagger||b| \right\|_\infty.$$

Using a demonstration similar to that of (9), this expression simplifies to

$$K_c(L, A, b) = \left\||D_{L^T x}^{-1}|(|V| vec(|A|) + |W||b|)\right\|_\infty, \tag{13}$$

and when $m = n$ (case of a linear system $Ax = b$), we obtain

$$K_c(L, A, b) = \left\||D_{L^T x}^{-1}||L^T A^{-1}|(|A||x| + |b|)\right\|_\infty, \tag{14}$$

which is the condition number given in [18] when $L = I$.

4. Conclusion.

We proved that working on the dual space enables us to derive condition numbers and we applied this to the case where perturbations on data are measured componentwise. By using this method, we obtained formulas for the conditioning of a linear functional of the linear least squares solution for which we provided an exact expression when the perturbations of the solution are measured using the infinity or a componentwise norm and an upper bound when using the Euclidean norm. We also gave the corresponding expressions for the special case of consistent linear systems.

References

[1] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, 3 edition, 1999.

[2] M. Arioli, M. Baboulin, and S. Gratton. A partial condition number for linear least-squares problems. *SIAM J. Matrix Anal. and Appl.*, 29(2):413–433, 2007.

[3] M. Arioli, J. W. Demmel, and I. S. Duff. Solving sparse linear systems with sparse backward error. *SIAM J. Matrix Anal. and Appl.*, 10(2):165–190, 1989.

[4] M. Arioli, I. S. Duff, and P. P. M de Rijk. On the augmented system approach to sparse least squares problems. *Numerische Mathematik*, 55:667–684, 1989.

[5] F. L. Bauer. Genauigkeitsfragen bei der Lösung linearer Gleichungssysteme. *Z. Angew. Math. Mech.*, 46(7):409–421, 1966.

[6] Å. Björck. Component-wise perturbation analysis and error bounds for linear least squares solutions. *BIT Numerical Mathematics*, 31:238–244, 1991.

[7] Å. Björck. *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, 1996.

[8] S. Chandrasekaran and I. C. F. Ipsen. On the sensitivity of solution components in linear systems of equations. *SIAM J. Matrix Anal. and Appl.*, 16(1):93–112, 1995.

[9] F. Cucker, H. Diao, and Y. Wei. On mixed and componentwise condition numbers for moore-penrose inverse and linear least squares problems. *Mathematics of Computation*, 76(258):947–963, 2007.

[10] J. Demmel, Y. Hida, X. S. Li, and E. J. Riedy. Extra-precise iterative refinement for overdetermined least squares problems. Technical Report EECS-2007-77, UC Berkeley, 2007. Also LAPACK Working Note 188.

[11] J. Dongarra, J. Du Croz, I. Duff, and S. Hammarling. A set of Level 3 Basic Linear Algebra Subprograms. *ACM Transactions on Mathematical Software*, 16:1–17, 1990.

[12] A. J. Geurts. A contribution to the theory of condition. *Numerische Mathematik*, 39:85–96, 1982.

[13] I. Gohberg and I. Koltracht. Mixed, componentwise, and structured condition numbers. *SIAM J. Matrix Anal. and Appl.*, 14(3):688–704, 1993.

[14] G. H. Golub and C. F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, 1996. Third edition.

[15] A. Graham. *Kronecker products and matrix calculus with application*. Wiley, New York, 1981.

[16] S. Gratton. On the condition number of linear least squares problems in a weighted Frobenius norm. *BIT Numerical Mathematics*, 36(3):523–530, 1996.

[17] J. F. Grcar. Adjoint formulas for condition numbers applied to linear and indefinite least squares. Technical Report LBNL-55221, Lawrence Berkeley National Laboratory, 2004.

[18] D. J. Higham and N. J. Higham. Componentwise perturbation theory for linear systems with multiple right-hand sides. *Linear Algebra and its Applications*, 174:111–129, 1992.

[19] N. J. Higham. Computing error bounds for regression problems. *Statistical Analysis of Measurement Error Models and Applications, Contemporary Mathematics 112 (Philip J. Brown and Wayne A. Fuller, eds.)*, pages 195–208, 1990. American Mathematical Society, Providence, RI, USA.

[20] N. J. Higham. A survey of componentwise perturbation theory in numerical linear algebra. *In W. Gautschi editor, Mathematics of Computation 1943-1993: A Half Century of Computational Mathematics, volume 48 of Proceedings of Symposia in Applied Mathematics*, pages 49–77, 1994. American Mathematical Society, Providence, RI, USA.

[21] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, 2 edition, 2002.
[22] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
[23] C. S. Kenney, A. J. Laub, and M. S. Reese. Statistical condition estimation for linear least squares. *SIAM J. Matrix Anal. and Appl.*, 19(4):906–923, 1998.
[24] P. Lancaster and M. Tismenetsky. *The theory of matrices*. Academic Press Inc., 1985. Second edition with applications.
[25] W. Oettli and W. Prager. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. *Numerische Mathematik*, 6:405–409, 1964.
[26] J. Rice. A theory of condition. *SIAM J. Numerical Analysis*, 3:287–310, 1966.
[27] J. Rohn. New condition numbers for matrices and linear systems. *Computing*, 41(1-2):167–169, 1989.
[28] S. M. Rump. Structured perturbations Part II: componentwise distances. *SIAM J. Matrix Anal. and Appl.*, 25(1):31–56, 2003.
[29] R. D. Skeel. Scaling for numerical stability in Gaussian elimination. *J. Assoc. Comput. Mach.*, 26(3):494–526, 1979.
[30] G. W. Stewart and Jiguang Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1991.

Marc Baboulin
CMUC, Department of Mathematics, University of Coimbra, Coimbra, Portugal
E-mail address: baboulin@mat.uc.pt

Serge Gratton
Centre National d'Etudes Spatiales, Toulouse, France
E-mail address: serge.gratton@cnes.fr