Numer. Math. 62, 17–33 (1992)

Numerische Mathematik

The structured sensitivity of Vandermonde-like systems

Sven G. Bartels* and Desmond J. Higham

Department of Mathematics and Computer Science, University of Dundee, Dundee, DD1 4HN, Scotland

Received October 15, 1991

Summary. We consider a general class of structured matrices that includes (possibly confluent) Vandermonde and Vandermonde-like matrices. Here the entries in the matrix depend nonlinearly upon a vector of parameters. We define condition numbers that measure the componentwise sensitivity of the associated primal and dual solutions to small componentwise perturbations in the parameters and in the right-hand side. Convenient expressions are derived for the infinity norm based condition numbers, and order-of-magnitude estimates are given for condition numbers defined in terms of a general vector norm. We then discuss the computation of the corresponding backward errors. After linearising the constraints, we derive an exact expression for the infinity norm dual backward error and show that the corresponding primal backward error is given by the minimum infinity-norm solution of an underdetermined linear system. Exact componentwise condition numbers are also derived for matrix inversion and the least squares problem, and the linearised least squares backward error is characterised.

Mathematics Subject Classification (1991): 65F99

1 Introduction

Many applications generate linear systems of the form $Ax=b$, $A\in\mathbb{R}^{N\times N}$, $b\in\mathbb{R}^N$, where the matrix A has a special structure. In such cases the data for the problem involves a set of parameters that determines A. For example, a symmetric Toeplitz matrix is defined by the values $\{c_0, c_1, \ldots, c_{N-1}\}$ that appear along the constant diagonals. With such a structured problem, any errors in the data produce structured errors in A, and hence it is pertinent to define a condition number with respect to a suitably restricted class of perturbations. Similarly,

* *Current address:* Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, W-7500 Karlsruhe, Federal Republic of Germany

Offprint requests to: D.J. Higham

given an approximate solution $\tilde{x}$, it is reasonable to define the backward error as the smallest perturbation to the original data that causes $\tilde{x}$ to become an exact solution.

Structured condition numbers and backward errors that are relevant to the case where A depends linearly upon a set of parameters were examined in [9]. Such linear dependence arises, for example, with symmetric, Toeplitz, circulant, Hankel and Hamiltonian matrices. In this work we consider the case where the matrix is generated by a set of real-valued functions $\{p_{ij}\}_{i,j=0}^{n}$. Given a set of real points $a_0 \leqq a_1 \leqq \dots \leqq a_n$, we define the matrix $V[a]$ by

$$(V[a])_{ij} = p_{ij}[a_j]. \tag{1.1}$$

We have in mind the case where $V[a]$ is a (possibly confluent) Vandermonde or Vandermonde-like matrix [12]. Here p_{ij} is a polynomial, and in the non-confluent case p_{ij} is independent of j. However, our results concerning condition numbers hold whenever the p_{ij} have continuous second derivatives in neighbourhoods of the points a_j, and the matrix $V[a]$ is nonsingular. Associated with $V[a]$ are the primal and dual problems

$$V[a]\,x = b, \quad \text{(PRIMAL)} \tag{1.2}$$

$$V[a]^{\mathrm{T}} x = b, \quad \text{(DUAL)}. \tag{1.3}$$

The primal problem arises in statistics with b representing the moments of a discrete random variable. The dual problem frequently occurs in approximation theory in the context of interpolation.

In examining the sensitivity of (1.2) and (1.3) we must decide how perturbations Δa, Δb and Δx to the data a, b and the solution x, respectively, are to be measured. We will adopt the following general componentwise measures. Choosing nonnegative tolerance vectors $\alpha, \beta, \xi \in \mathbb{R}^{n+1}$, we form

$$\|\widehat{\Delta a}\|_v, \quad \|\widehat{\Delta b}\|_v, \quad \|\widehat{\Delta x}\|_v, \tag{1.4}$$

where

$$\widehat{\Delta a}_i = \Delta a_i/\alpha_i, \quad \widehat{\Delta b}_i = \Delta b_i/\beta_i, \quad \widehat{\Delta x}_i = \Delta x_i/\xi_i, \tag{1.5}$$

and $\|\cdot\|_v$ represents any vector norm. Here, and throughout, we use the division-by-zero convention that $\varepsilon/0$ is zero if $\varepsilon = 0$ and infinity otherwise. Note that choosing $\alpha_i = \|a\|_v$, $\beta_i = \|b\|_v$ and $\xi_i = \|x\|_v$ corresponds to a traditional normwise relative measure of the perturbations, while the choice $\alpha_i = |a_i|$, $\beta_i = |b_i|$ and $\xi_i = |x_i|$ gives a componentwise relative measure.

In subsequent sections we use the measure (1.4) to define structured condition numbers and backward errors. Our aim is then to find convenient characterisations or bounds for these quantities, and to investigate their computability. The next section looks at structured condition numbers for the primal and dual linear systems. An exact expression for the ∞-norm case is derived, and in other cases we obtain an upper bound that is within a factor two of equality. We also look at the case where the solution to the dual problem is regarded as the interpolating function, $\Sigma_i x_i p_i$, and is measured with a corresponding functional norm. We show that the resulting condition number can be approximated very conveniently. Backward errors for (1.2) and (1.3) are considered in Sect. 3. We find that in order to derive useful results it is necessary to linearise the constraints and settle for an approximate backward error. In Sect. 4 we compare the structured condition numbers and backward errors with their unstructured counterparts. Exact condition number expressions for inverting $V[a]$ and $V[a]^T$, with a certain choice of matrix norm, are furnished in Sect. 5. The least squares analogues of (1.2) and (1.3) are considered in Sect. 6. We obtain exact expressions for the primal and dual condition numbers when $\|\cdot\|_v = \|\cdot\|_\infty$, and sharp bounds for general $\|\cdot\|_v$. We also show how to compute the linearised backward error in the case of the ∞-norm.

We conclude this section by mentioning some related work. The idea of measuring perturbations in a componentwise, rather than a normwise sense has been investigated by Oettli and Prager [16], Skeel [18] and more recently by Arioli et al. [1], Rohn [17], and, for the least squares problem, by Björck [3] and N.J. Higham [14]. In these papers no special structure in the matrix A is assumed, other than sparsity. Symmetric normwise backward errors were investigated by Bunch et al. [5], and componentwise condition numbers and backward errors for the case of general linear dependence on a set of parameters were examined by Higham and Higham [9, 10, Sect. 5]. Gohberg and Koltracht [6] gave a formal definition of the condition number of a structured problem – in our case this corresponds to setting $\|\cdot\|_v = \|\cdot\|_\infty$ and using $\alpha = |a|$, $\beta = |b|$ and $\xi = |x|$ in (1.4). Bounds on the structured condition number for Cauchy matrix inversion were given in [6]. The structured componentwise sensitivity of primal and dual Vandermonde systems was characterised by N.J. Higham in Sect. 4 of [11]. Those results are special cases of our results in Sect. 2. When discussing the cost of evaluating the expressions that we derive in the forthcoming sections, we will assume that $V[a]$ is a confluent Vandermonde-like matrix and we will make use of the fact that the corresponding primal and dual problems (1.2) and (1.3) can be solved in $O(n^2)$ operations by the algorithms of N.J. Higham [12], which include as special cases the fast Vandermonde system solvers of Björck and Pereyra [4]. We will also make reference to the fact that for any matrix B, estimates of $\|B\|_1$ or $\|B\|_\infty$ can be obtained at the cost of forming a small number of matrix-vector products of the form Bz and $B^T z$ [2, 8, 13, 19].

2 Primal and dual condition numbers

How sensitive is the solution x in (1.2) and (1.3) to small changes in the data a and b? The answer, at least in the limit as the perturbation size tends to zero, is given by the condition number. Considering first the dual problem, we define the componentwise structured condition number to be

$$\text{(2.1)} \qquad \operatorname{cond}_v^{\text{dual}}(a,b) := \lim_{\varepsilon \to 0} \sup_{\|\widehat{\Delta a}\|_v \le \varepsilon,\ \|\widehat{\Delta b}\|_v \le \varepsilon} \cdot \left\{ \frac{\|\widehat{\Delta x}\|_v}{\varepsilon} :\ V[a+\Delta a]^T (x+\Delta x) = b + \Delta b \right\}.$$

The term $\|\widehat{\Delta x}\|_v/\varepsilon$ inside the curly brackets gives the change in the solution x relative to the change in the data a, b, and the overall condition number is found by taking the limit as $\varepsilon \to 0$ of the worst case relative change in x caused by data perturbations of size ε.

Initially, we will assume that the tolerances α, β and ξ have nonzero elements. Zero tolerances will be discussed later. Using $\|\widehat{\Delta a}\|_v \leqq \varepsilon$ and $\|\widehat{\Delta b}\|_v \leqq \varepsilon$, and replacing $p_{ij}[a_j + \Delta a_j]$ by $p_{ij}[a_j] + \Delta a_j p'_{ij}[a_j] + O(\varepsilon^2)$, we find

$$(2.2) \qquad V[a+\Delta a] = V[a] + V'[a]\,\mathrm{diag}(\Delta a) + O(\varepsilon^2),$$

where $(V'[a])_{ij} := \frac{d}{d a_j}(V[a])_{ij} = p'_{ij}[a_j]$. Hence the constraint $V[a+\Delta a]^{\mathrm{T}}(x + \Delta x) = b + \Delta b$ in (2.1) becomes

$$(2.3) \qquad V[a]^{\mathrm{T}} \Delta x = \Delta b - \mathrm{diag}(\Delta a)\, V'[a]^{\mathrm{T}} x + O(\varepsilon^2).$$

To simplify the notation, we will let $z = V'[a]^{\mathrm{T}} x$. Note that in the case of polynomial interpolation, z_i denotes the derivative of the interpolating polynomial at the point a_i. We may then write (2.3) as

$$(2.4) \qquad \Delta x = V[a]^{-\mathrm{T}}(\Delta b - \mathrm{diag}(z)\,\Delta a) + O(\varepsilon^2),$$

where $V[a]^{-\mathrm{T}} := (V[a]^{-1})^{\mathrm{T}}$. Writing D_α, D_β and D_ξ for diag(α), diag(β) and diag(ξ) respectively, we may write (2.4) in terms of the scaled perturbations $\widehat{\Delta a}$, $\widehat{\Delta b}$ and $\widehat{\Delta x}$ in (1.5):

$$(2.5) \qquad \widehat{\Delta x} = D_\xi^{-1}\, V[a]^{-\mathrm{T}}(D_\beta \widehat{\Delta b} - \mathrm{diag}(z)\, D_\alpha \widehat{\Delta a}) + O(\varepsilon^2).$$

Taking norms, with $\|\widehat{\Delta a}\|_v \leqq \varepsilon$ and $\|\widehat{\Delta b}\|_v \leqq \varepsilon$, it follows that

$$\|\widehat{\Delta x}\|_v \leqq \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} D_\beta\|_v\, \varepsilon + \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} \mathrm{diag}(z)\, D_\alpha\|_v\, \varepsilon + O(\varepsilon^2),$$

and hence

$$(2.6) \qquad \mathrm{cond}_v^{\mathrm{dual}}(a,b) \leqq \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} D_\beta\|_v + \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} \mathrm{diag}(z)\, D_\alpha\|_v,$$

where the matrix norm $\|\cdot\|_v$ is subordinate to the corresponding vector norm. Note that $\widehat{\Delta b}$ and $\widehat{\Delta a}$ can be chosen in (2.5) so that

$$\|D_\xi^{-1}\, V[a]^{-\mathrm{T}} D_\beta \widehat{\Delta b}\|_v = \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} D_\beta\|_v\, \varepsilon,$$
$$\|D_\xi^{-1}\, V[a]^{-\mathrm{T}} \mathrm{diag}(z)\, D_\alpha \widehat{\Delta a}\|_v = \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} \mathrm{diag}(z)\, D_\alpha\|_v\, \varepsilon,$$

and so it follows that

$$\|D_\xi^{-1}\, V[a]^{-\mathrm{T}} D_\beta\|_v + \|D_\xi^{-1}\, V[a]^{-\mathrm{T}} \mathrm{diag}(z)\, D_\alpha\|_v \leqq 2\, \mathrm{cond}_v^{\mathrm{dual}}(a,b).$$

When the infinity norm is used, we can make use of the following lemma.

Lemma 2.1. *For any* $A, B \in \mathbb{R}^{N \times N}$

$$(2.7) \qquad \sup_{\|Ac\|_\infty \leqq \varepsilon,\ \|d\|_\infty \leqq \varepsilon} \{\|Ac + Bd\|_\infty\} = \||A| + |B|\|_\infty\, \varepsilon.$$

Proof. Let a_i^{T} and b_i^{T} denote the ith rows of A and B, respectively, and let e denote the vector of ones. Then

$$\begin{aligned}\sup_{\|c\|_\infty \leqq \varepsilon,\ \|d\|_\infty \leqq \varepsilon} \{\|Ac+Bd\|_\infty\} &= \sup_{\|c\|_\infty \leqq 1,\ \|d\|_\infty \leqq 1} \{\max_i |a_i^{\mathrm{T}} c + b_i^{\mathrm{T}} d|\}\, \varepsilon \\ &= \max_i (|a_i|^{\mathrm{T}} e + |b_i|^{\mathrm{T}} e)\, \varepsilon \\ &= \||A|+|B|\|_\infty\, \varepsilon. \quad \square\end{aligned}$$

Using this lemma in (2.5) we obtain an exact expression for the ∞-norm condition number,

$$\begin{aligned}\operatorname{cond}_\infty^{\mathrm{dual}}(a,b) &= \|D_\xi^{-1} V[a]^{-\mathrm{T}} \operatorname{diag}(\beta + D_\alpha |z|)\|_\infty, \\ &= \|D_\xi^{-1} |V[a]^{-\mathrm{T}}| (\beta + D_\alpha |z|)\|_\infty. \end{aligned} \tag{2.8}$$

We now discuss the case where some tolerance values are zero. If some $\alpha_i = 0$ or $\beta_i = 0$ then, using the division-by-zero convention, the definition (2.1) is still valid. In this case, in order to keep $\|\widehat{\Delta a}\|_v \leqq \varepsilon$ and $\|\widehat{\Delta b}\|_v \leqq \varepsilon$, the data values a_i and b_i for which the tolerances are zero must not be perturbed, that is

$$\alpha_i = 0 \Rightarrow \Delta a_i = 0 \quad \text{and} \quad \beta_i = 0 \Rightarrow \Delta b_i = 0.$$

It is also clear that (2.6) and (2.8) remain valid in this case. Suppose now that one of the solution tolerances ξ_i is zero. It is then possible for the condition number to be infinite: if, for all $\varepsilon > 0$, there exist feasible perturbations $\|\widehat{\Delta a}\|_v \leqq \varepsilon$ and $\|\widehat{\Delta b}\|_v \leqq \varepsilon$ such that $\Delta x_i \neq 0$, then $\operatorname{cond}_v^{\mathrm{dual}}(a,b) = \infty$. It is straightforward to confirm that the right-hand sides of (2.6) and (2.8) become infinite under exactly the same circumstances, and hence the results are still meaningful. In a similar manner, all the results in this paper can be interpreted sensibly when one or more tolerance values are zero.

We thus have the following theorem.

Theorem 2.1. *In the notation above and for any vector norm, the structured componentwise condition number for the dual problem satisfies*

$$\operatorname{cond}_v^{\mathrm{dual}}(a,b) \leqq \|D_\xi^{-1} V[a]^{-\mathrm{T}} D_\beta\|_v + \|D_\xi^{-1} V[a]^{-\mathrm{T}} \operatorname{diag}(z) D_\alpha\|_v, \tag{2.9}$$

and this upper bound is not more than twice $\operatorname{cond}_v^{\mathrm{dual}}(a,b)$. *For the infinity vector norm,*

$$\operatorname{cond}_\infty^{\mathrm{dual}}(a,b) = \|D_\xi^{-1} V[a]^{-\mathrm{T}} \operatorname{diag}(\beta + D_\alpha |z|)\|_\infty. \tag{2.10}$$

The corresponding structured componentwise condition number for the primal problem is

$$\operatorname{cond}_v^{\mathrm{primal}}(a,b) := \lim_{\varepsilon \to 0} \sup_{\|\widehat{\Delta a}\|_v \leqq \varepsilon,\ \|\widehat{\Delta b}\|_v \leqq \varepsilon} \left\{ \frac{\|\widehat{\Delta x}\|_v}{\varepsilon} : V[a+\Delta a]\,(x+\Delta x) = b + \Delta b \right\}. \tag{2.11}$$

Expanding the constraint $V[a+\Delta a](x+\Delta x)=b+\Delta b$ in the same manner as in the dual case, we find

$$V[a]\,\Delta x=\Delta b-V'[a]\,\mathrm{diag}(x)\,\Delta a+O(\varepsilon^2).$$

Defining $\bar{V}[a,x]=V'[a]\,\mathrm{diag}(x)$, and converting to the scaled perturbations, we have

$$\widehat{\Delta x}=D_\xi^{-1}\,V[a]^{-1}(D_\beta\,\widehat{\Delta b}-\bar{V}[a,x]\,D_\alpha\,\widehat{\Delta a})+O(\varepsilon^2). \tag{2.12}$$

As for the dual case, we may take norms in (2.12) to get an upper bound on $\mathrm{cond}_v^{\mathrm{primal}}(a,b)$, and Lemma 2.1 can be used to give an exact expression for $\mathrm{cond}_\infty^{\mathrm{primal}}(a,b)$.

The results are collected in the following theorem.

Theorem 2.2. *In the notation above and for any vector norm, the structured componentwise condition number for the primal problem satisfies*

$$\mathrm{cond}_v^{\mathrm{primal}}(a,b)\leqq\|D_\xi^{-1}\,V[a]^{-1}D_\beta\|_v+\|D_\xi^{-1}\,V[a]^{-1}\,\bar{V}[a,x]\,D_\alpha\|_v, \tag{2.13}$$

and this upper bound is not more than twice $\mathrm{cond}_v^{\mathrm{primal}}(a,b)$. *For the infinity vector norm,*

$$\mathrm{cond}_\infty^{\mathrm{primal}}(a,b)=\|\,|D_\xi^{-1}\,V[a]^{-1}D_\beta|+|D_\xi^{-1}\,V[a]^{-1}\,\bar{V}[a,x]\,D_\alpha|\,\|_\infty. \tag{2.14}$$

As we mentioned in Sect. 1, these results can be thought of as extensions to those in Sect. 4 of [11]. N.J. Higham considered the standard Vandermonde matrix for which $p_{ij}(a_j)=a_j^i$ in (1.1), and looked at a componentwise relative perturbation using $\|\cdot\|_\infty$, with perturbations to the matrix and the right-hand side considered separately. In our notation this corresponds to $\alpha=0$, $\beta=|b|$, $\xi_i=\|x\|_\infty$ and $\alpha=|a|$, $\beta=0$, $\xi_i=\|x\|_\infty$. Our expressions for $\mathrm{cond}_\infty^{\mathrm{primal}}(a,b)$ and $\mathrm{cond}_\infty^{\mathrm{dual}}(a,b)$ then reduce to those in [11].

In order to evaluate the bounds (2.9) and (2.13), or the expressions (2.10) and (2.14), one must, in effect, invert $V[a]$ or $V[a]^{\mathrm{T}}$. Since the primal and dual problems (1.2) and (1.3) can be solved in $O(n^2)$ operations, inversion via $n+1$ linear systems is an $O(n^3)$ process. Traub [20] has shown that a Vandermonde matrix can be inverted in $O(n^2)$ operations, but this algorithm has questionable stability properties; see Sect. 3 of [11]. A viable alternative for $\|\cdot\|_v=\|\cdot\|_\infty$ or $\|\cdot\|_v=\|\cdot\|_1$ is to use a matrix norm estimator, such as ACM TOMS Algorithm 674 [13]. For any matrix B, this algorithm computes a lower bound on $\|B\|_\infty$ or $\|B\|_1$ that is almost always within a factor ten of the true norm, at a cost of no more than ten matrix-vector products of the form Bz and $B^{\mathrm{T}}z$. With this approach the condition numbers in Theorems 2.1 and 2.2 can be approximated in $O(n^2)$ operations (ignoring the cost of forming $V'[a]$). For example, in (2.9) we have $B=D_\xi^{-1}\,V[a]^{-\mathrm{T}}D_\beta$ and $B=D_\xi^{-1}\,V[a]^{-\mathrm{T}}\,\mathrm{diag}(z)\,D_\alpha$ and the cost of forming Bz or $B^{\mathrm{T}}z$ is dominated by the cost of solving a dual or primal problem. For the primal condition numbers in Theorem 2.2, the presence of the $\bar{V}[a,x]$ factor increases the expense of Bz and $B^{\mathrm{T}}z$, but the overall cost remains $O(n^2)$.

We consider now the special case of non-confluent Vandermonde-like matrices – here p_{ij} is a polynomial which is independent of j, that is $p_{ij}\equiv p_i$. Suppose that the solution to the dual problem (1.3) represents the interpolating polynomial $\sum_{i=0}^{n} x_i p_i$. It may then be appropriate to measure the solution using a functional norm, such as an L_p norm;

$$\|x\|_{L_p} = \left(\int_{a_0}^{a_n} \left| \sum_{i=0}^{n} x_i p_i(y) \right|^p dy \right)^{1/p}, \qquad 1 \leq p < \infty. \tag{2.15}$$

Note that $\|\cdot\|_{L_p}$ is a vector norm on $\mathbb{R}^{n+1}$.

An associated condition number is then

$$\operatorname{cond}^{\text{dual}}_{v, L_p}(a, b) := \lim_{\varepsilon \to 0} \sup_{\|\widehat{\Delta a}\|_v \leq \varepsilon,\ \|\widehat{\Delta b}\|_v \leq \varepsilon} \cdot \left\{ \frac{\|\Delta x\|_{L_p}}{\varepsilon} : V[a+\Delta a]^{\mathrm{T}}(x+\Delta x) = b + \Delta b \right\}. \tag{2.16}$$

For simplicity, we restrict our attention to the $\operatorname{cond}^{\text{dual}}_{\infty, L_p}(a, b)$ case. From (2.4), we see that the key quantity is

$$\sup_{|\Delta a| \leq \varepsilon\alpha,\ |\Delta b| \leq \varepsilon\beta} \left(\int_{a_0}^{a_n} \left| \sum_{i=0}^{n} \left(\sum_{j=0}^{n} (V[a]^{-\mathrm{T}})_{ij} (\Delta b_j - z_j \Delta a_j) \right) p_i(y) \right|^p dy \right)^{1/p},$$

which may be rearranged as

$$\sup_{|\Delta a| \leq \varepsilon\alpha,\ |\Delta b| \leq \varepsilon\beta} \left(\int_{a_0}^{a_n} \left| \sum_{j=0}^{n} \left(\sum_{i=0}^{n} (V[a]^{-\mathrm{T}})_{ij} p_i(y) \right) (\Delta b_j - z_j \Delta a_j) \right|^p dy \right)^{1/p}. \tag{2.17}$$

If we approximate the definite integral in (2.17) using a quadrature rule with support abscissae $a_0, a_1, \ldots, a_n$, then the expression simplifies dramatically. Suppose that the quadrature rule has the form

$$\int_{a_0}^{a_n} f(y)\, dy \approx \sum_{k=0}^{n} w_k f(a_k).$$

Then the approximation to (2.17) is

$$\sup_{|\Delta a| \leq \varepsilon\alpha,\ |\Delta b| \leq \varepsilon\beta} \left(\sum_{k=0}^{n} w_k \left| \sum_{j=0}^{n} \left(\sum_{i=0}^{n} (V[a]^{-\mathrm{T}})_{ij} p_i(a_k) \right) (\Delta b_j - z_j \Delta a_j) \right|^p \right)^{1/p},$$

and since $p_i(a_k) = (V[a])_{ik}$, this reduces to

$$\sup_{|\Delta a| \leq \varepsilon\alpha,\ |\Delta b| \leq \varepsilon\beta} \left(\sum_{k=0}^{n} w_k |\Delta b_k - z_k \Delta a_k|^p \right)^{1/p} = \left(\sum_{k=0}^{n} w_k (\beta_k + |z_k|\, \alpha_k)^p \right)^{1/p} \varepsilon.$$

Hence $\left(\sum_{k=0}^{n} w_k (\beta_k + |z_k|\, \alpha_k)^p \right)^{1/p}$ provides an extremely inexpensive approximation to $\operatorname{cond}^{\text{dual}}_{\infty, L_p}(a, b)$.

There are two possible drawbacks with this approach. One is that the quadrature method is restricted to the grid $\{a_i\}_{i=0}^n$, which may be very coarse or irregular. The other is that the integrand in (2.17) is generally non-differentiable, and for both these reasons the integral approximations might not be very accurate. In order to avoid the non-differentiability, we suggest using the L_2 measure.

3 Primal and dual backward errors

Given an approximate solution $\tilde{x}$ to (1.2) or (1.3), how close is the nearest system of the same structure for which $\tilde{x}$ is an exact solution? Considering the dual problem first, this question leads to the definition of the componentwise structured backward error

$$\text{(3.1)} \quad \mathrm{be}_v^{\mathrm{dual}}(a, b, \tilde{x}) := \inf\{\varepsilon : V[a+\Delta a]^{\mathrm{T}} \tilde{x} = b + \Delta b,\ \|\widehat{\Delta a}\|_v \leqq \varepsilon,\ \|\widehat{\Delta b}\|_v \leqq \varepsilon\}.$$

Note that if there are no feasible perturbations Δa and Δb such that $\|\widehat{\Delta a}\|_v$ and $\|\widehat{\Delta b}\|_v$ are finite, then the backward error is regarded as infinite.

The dual case is simpler than the corresponding primal case because each perturbation appears in only one constraint equation:

$$\text{(3.2)} \qquad \sum_{j=0}^{n} p_{ji}[a_i + \Delta a_i]\, \tilde{x}_j = b_i + \Delta b_i, \qquad 0 \leqq i \leqq n.$$

If we use $\|\cdot\|_v = \|\cdot\|_\infty$ in (3.1), then we can decouple the problem into $n+1$ independent subproblems of the form

$$\text{(3.3)} \qquad \inf\left\{\max\left\{\frac{|\Delta a_i|}{\alpha_i}, \frac{|\Delta b_i|}{\beta_i}\right\} : \sum_{j=0}^{n} p_{ji}[a_i + \Delta a_i]\, \tilde{x}_j = b_i + \Delta b_i\right\}.$$

Note that if $\alpha_i = 0$ in (3.3) then Δa_i is forced to be zero, and similarly $\beta_i = 0$ forces $\Delta b_i = 0$. In the case where p_{ji} are polynomials, (3.3) has the following interpretation. Given the polynomial $P_i(y) = \sum_{j=0}^{n} \tilde{x}_j p_{ji}(y)$ and the point (a_i, b_i), find the smallest perturbation, measured as $\|[\Delta a_i/\alpha_i, \Delta b_i/\beta_i]^{\mathrm{T}}\|_\infty$, such that $P_i(y)$ interpolates $(a_i + \Delta a_i,\ b_i + \Delta b_i)$. It will usually be the case that the optimum perturbation satisfies $|\Delta a_i|/\alpha_i = |\Delta b_i|/\beta_i$, since if not, the larger of the two values can normally be decreased at the expense of (possibly) increasing the other. An exception can occur when $|\Delta a_i|/\alpha_i < |\Delta b_i|/\beta_i$ and $a_i + \Delta a_i$ is a local maximum or minimum of $P_i(y)$. Hence, the optimum point $(a_i + \Delta a_i, b_i + \Delta b_i)$ is either

(1) An intersection point of $P_i(y)$ and one of the lines $\Delta a_i\, \beta_i = \pm \Delta b_i\, \alpha_i$.
(2) A local maximum or minimum of $P_i(y)$.

Hence $\mathrm{be}_\infty^{\mathrm{dual}}(a, b, \tilde{x})$ can be computed by finding, for $i = 0, 1, \ldots, n$, the roots of two polynomials of degree $n+1$ and one polynomial of degree n. In the case of a non-confluent Vandermonde-like matrix, $p_{ji}(y)$ is independent of i, and so step (2) need only be done once.

Since polynomial root-finding is an expensive process, especially for large n, an attractive alternative is to linearize the constraint in (3.2). We will denote the corresponding linearised backward error approximation to $\mathrm{be}_v^{\mathrm{dual}}(a, b, \tilde{x})$ by $\mathrm{linbe}_v^{\mathrm{dual}}(a, b, \tilde{x})$. The linearised subproblem for the $\|\cdot\|_\infty$ backward error may be written

$$\text{(3.4)} \qquad \inf\left\{\max\left\{\frac{|\Delta a_i|}{\alpha_i}, \frac{|\Delta b_i|}{\beta_i}\right\}: \Delta a_i \tilde{z}_i - \Delta b_i = r_i\right\},$$

where $\tilde{z} := V'[a]^T \tilde{x}$ and the residual $r := b - V[a]^T \tilde{x}$. We may re-write the constraint as

$$\left[\frac{\Delta a_i}{\alpha_i}, \frac{\Delta b_i}{\beta_i}\right]\begin{bmatrix}\tilde{z}_i \alpha_i \\ -\beta_i\end{bmatrix} = r_i.$$

The Hölder inequality (see, for example, [15]) then says that for any Hölder p- and q-norms with $p, q \geqq 1$ and $1/p + 1/q = 1$,

$$\text{(3.5)} \qquad \left\|\left[\frac{\Delta a_i}{\alpha_i}, \frac{\Delta b_i}{\beta_i}\right]^T\right\|_p \geqq \frac{|r_i|}{\|[\tilde{z}_i \alpha_i, \beta_i]^T\|_q},$$

with equality being attainable for certain perturbations. Using $p = \infty$, we find that

$$\text{(3.6)} \qquad \mathrm{linbe}_\infty^{\mathrm{dual}}(a, b, \tilde{x}) = \max_i \left\{\frac{|r_i|}{(\beta + D_\alpha |\tilde{z}|)_i}\right\}. \qquad \text{(3.6)}$$

The corresponding optimal perturbations are given by

$$\Delta a_i^* = \frac{\mathrm{sign}(\tilde{z}_i)\, r_i \alpha_i}{|\tilde{z}_i| \alpha_i + \beta_i}, \qquad \Delta b_i^* = \frac{-r_i \beta_i}{|\tilde{z}_i| \alpha_i + \beta_i}.$$

Note that a genuine upper bound on $\mathrm{be}_\infty^{\mathrm{dual}}(a, b, \tilde{x})$ can be constructed from Δa^* by making a further perturbation to the right-hand side. Letting

$$\Delta b_i^{**} = \sum_{j=0}^{n} p_{ji}[a_i + \Delta a_i^*]\, \tilde{x}_j - b_i,$$

$\{\Delta a^*, \Delta b^{**}\}$ is clearly a feasible perturbation, and hence

$$\text{(3.7)} \qquad \mathrm{be}_\infty^{\mathrm{dual}}(a, b, \tilde{x}) \leqq \max\{\|\widehat{\Delta a^*}\|_\infty, \|\widehat{\Delta b^{**}}\|_\infty\}.$$

We mention that in the worst case this bound can be arbitrarily poor. In particular, if $\beta_i = 0$ and $\Delta b_i^{**} \neq 0$, then the right-hand side of (3.7) is infinite, even though the true backward error could be finite.

For a general Hölder p-norm, minimizing $\|[\Delta a_i/\alpha_i, \Delta b_i/\beta_i]^T\|_p$ for each i, using (3.5), leads to an upper bound for $\mathrm{linbe}_v^{\mathrm{dual}}(a, b, \tilde{x})$,

$$\text{(3.8)} \qquad \mathrm{linbe}_v^{\mathrm{dual}}(a, b, \tilde{x}) \leqq \left\|\left(\frac{r_i}{\|[\tilde{z}_i \alpha_i, \beta_i]^T\|_q}\right)\right\|_p.$$

As with the $\text{linbe}_\infty^{\text{dual}}(a, b, \tilde{x})$ case above, this can be converted into an upper bound for $\text{be}_v^{\text{dual}}(a, b, \tilde{x})$ by re-perturbing the right-hand side.

The analogous definition to (3.1) for the primal problem is

$$(3.9)\quad \text{be}_v^{\text{primal}}(a, b, \tilde{x}) := \inf\{\varepsilon \colon V[a+\Delta a]\,\tilde{x} = b + \Delta b, \|\widehat{\Delta a}\|_v \leqq \varepsilon, \|\widehat{\Delta b}\|_v \leqq \varepsilon\}.$$

We now have a nonlinear, non-convex optimization problem with $2n+3$ variables.

Moving directly to the linearised approximate backward error, the constraint becomes

$$(3.10)\qquad V'[a]\,\text{diag}(\tilde{x})\,\Delta a - \Delta b = r,$$

where r now denotes the primal residual, $r := b - V[a]\,\tilde{x}$. The constraint can be re-written

$$(3.11)\qquad [V'[a]\,\text{diag}(\tilde{x})\,D_\alpha, \; -D_\beta] \begin{bmatrix} \widehat{\Delta a} \\ \widehat{\Delta b} \end{bmatrix} = r.$$

This is an underdetermined system of $n+1$ linear equations in the $2(n+1)$ unknowns. The problem of finding the minimum-norm solution to (3.11) can be transformed into the problem of finding the minimum-norm residual to an overdetermined system (see Sect. 2 of [10] for more details). Standard methods exist for such problems when the Hölder p-norm for $p = 1, 2, \infty$ is used. In the $\|\cdot\|_\infty$ case, this is equivalent to computing $\text{linbe}_\infty^{\text{primal}}(a, b, \tilde{x})$. In the $\|\cdot\|_2$ and $\|\cdot\|_1$ cases, the norm of the minimum-norm solution to (3.11) is clearly an upper bound for the corresponding $\text{linbe}_v^{\text{primal}}(a, b, \tilde{x})$. As above, an upper bound for the actual structured backward error can be generated from the linearised backward error perturbations by altering Δb.

For the dual problem, the expressions (3.6), (3.7) and (3.8) can be computed in $O(n^2)$ operations. In the primal case, finding the minimum-norm solution to (3.11) is $O(n^3)$ for $p = 2$, and requires an iterative algorithm for $p = 1, \infty$.

4 Discussion

How do the structured condition numbers and backward errors derived in the last two sections compare with their unstructured counterparts? Since the structured definitions involve a restricted class of perturbations, intuitively, the condition number should decrease and the backward error should increase under the imposition of structure. However, in order to make this argument precise, we must measure the perturbations consistently.

An unstructured componentwise condition number for the problem $Ax = b$ can be defined as

$$(4.1)\quad \text{ucond}(A, b) := \lim_{\varepsilon \to 0} \sup_{|\Delta A| \leqq \varepsilon E, \, |\Delta b| \leqq \varepsilon \beta} \left\{ \frac{\|\widehat{\Delta x}\|_\infty}{\varepsilon} \colon (A + \Delta A)(x + \Delta x) = b + \Delta b \right\}.$$

Here $\widehat{\Delta x_i} := x_i/\xi_i$ (as usual) and E is a matrix of nonnegative tolerances. This condition number can be compared with the $\|\cdot\|_\infty$ structured condition number

Table 1. Condition number ratios: unstructured/structured

	max	min	mean
(1)	4.51	0.63	1.28
(2)	2.96	0.68	1.20
(3)	15705.40	145.52	1862.76

– in both cases we are allowed to perturb each component of the data by an amount ε, relative to the tolerance value. We can characterise (4.1) as

$$\text{ucond}(A, b) = \|D_\xi^{-1} |A^{-1}| (\beta + E|x|)\|_\infty . \tag{4.2}$$

In the case where $\xi_i = \|x\|_\infty$, this result is given in Eq. (3.5) of [9]. The extension to general ξ_i is straightforward. From the expansion (2.2) it follows that

$$D_\alpha |V'[a]^{\mathrm{T}}| \leqq E \Rightarrow \text{cond}_\infty^{\text{dual}}(a, b) \leqq \text{ucond}(V[a]^{\mathrm{T}}, b),$$
$$|V'[a]|\, D_\alpha \leqq E \Rightarrow \text{cond}_\infty^{\text{primal}}(a, b) \leqq \text{ucond}(V[a], b)$$

(assuming that the same β and ξ are used in the two cases). This result can also be deduced by comparing (2.10) and (2.14) with (4.2). The unstructured componentwise backward error for $Ax = b$

$$\text{ube}(A, b, \tilde{x}) := \inf\{\varepsilon : (A + \Delta A)\, \tilde{x} = b + \Delta b, |\Delta A| \leqq \varepsilon E, |\Delta b| \leqq \varepsilon \beta\},$$

has the characterisation [16]

$$\text{ube}(A, b, \tilde{x}) = \max_i \left\{ \frac{|r_i|}{(\beta + E|\tilde{x}|)_i} \right\}, \tag{4.3}$$

where $r := b - Ax$. It follows from (3.6) and (4.3) that

$$D_\alpha |V'[a]^{\mathrm{T}}| \leqq E \Rightarrow \text{linbe}_\infty^{\text{dual}}(a, b, \tilde{x}) \geqq \text{ube}(V[a]^{\mathrm{T}}, b, \tilde{x}),$$

$$|V'[a]| D_\alpha \leqq E \Rightarrow \text{linbe}_\infty^{\text{primal}}(a, b, \tilde{x}) \geqq \text{ube}(V[a], b, \tilde{x}).$$

Note, however, that with the "natural" componentwise relative tolerances $\alpha = |a|$ and $E = |V[a]|$ (primal) or $E = |V[a]^{\mathrm{T}}|$ (dual), it is not possible, in general, to predict whether the structured condition numbers or backward errors will be smaller or larger than the unstructured versions. For example, from (2.8) and (4.2) we see that the ratio of $\text{ucond}(V[a]^{\mathrm{T}}, b)$ to $\text{cond}_\infty^{\text{dual}}(a, b)$ depends largely on the relative sizes of β, $D_\alpha |z|$ and $E|x|$. As an illustration, we give some numerical results for the dual problem with the standard Vandermonde matrix $(V[a])_{ij} = a_j^i$, and with tolerances $\alpha = |a|$, $\beta = |b|$, $\xi = |x|$ and $E = |V[a]^{\mathrm{T}}|$. Three sets of data points $a \in \mathbb{R}^6$ were chosen:

(1) The Chebyshev polynomial zeros, $a_i = \cos((i + 0.5)\, \pi/6)$, $i = 0, 1, \ldots, 5$.
(2) Equally spaced points in $[-1, 1]$.
(3) Equally spaced points in $[1, 1.5]$.

For each a we generated 1000 right-hand sides b with elements from the Normal (0, 1) distribution, and computed the ratios $\text{ucond}(V[a]^{\mathrm{T}}, b)/\text{cond}_\infty^{\text{dual}}(a, b)$. The maximum, minimum and mean values of the ratios are given in Table 1. For sets (1) and (2) the ratios are typically greater than one, but occasionally take values less than one. For set (3) the ratios are much larger – here $|V[a]^T|\,|x| \gg |b|$ and cancellation in the product $z = V'[a]^T x$ causes $D_\alpha |z| \ll |V[a]^T|\,|x|$. These results emphasise that the usual, unstructured condition numbers can be inappropriate for predicting the effect of structured perturbations.

The simple form of the standard Vandermonde matrix allows us to determine a lower bound on $\text{ucond}(V[a]^T, b)/\text{cond}_\infty^{\text{dual}}(a, b)$ when $\alpha = |a|$ and $E = |V[a]^T|$. We have

$$(D_\alpha |V'[a]^T|)_{ij} = j(|V[a]^T|)_{ij} \leqq n(|V[a]^T|)_{ij},$$

from which it follows that

$$\frac{\text{ucond}(V[a]^T, b)}{\text{cond}_\infty^{\text{dual}}(a, b)} \geqq \frac{2}{n+1}.$$

A similar result holds in the primal case.

It is worth emphasising that although structured condition numbers are clearly of importance in the case where the input data contains errors, they are not necessarily relevant for investigating the accuracy of computed solutions. Numerical algorithms for structured linear systems do not generally guarantee to solve a nearby system with the same structure. In particular, we know of no Vandermonde algorithm with this property. For the same reason, whilst it may well be of interest to compute, or approximate, the structured backward error, this quantity will not necessarily be small.

Condition numbers and backward errors can also be used to provide forward error bounds. It is clear that the forward error $\|\tilde{x} - x\|_v$ can be approximately bounded by the product of the condition number and backward error, assuming that they are defined consistently. However, we can see no reason why the structured product should be preferred to the unstructured product, and numerical experiments in [2] revealed little difference between the two cases.

5 Matrix inversion condition numbers

Although it is rarely necessary to invert a matrix in practice, it is traditional to consider the sensitivity of this process, and to define an associated condition number. This quantity has the virtue of giving information about a matrix that does not depend on a particular right-hand side.

Using a vector norm $\|\cdot\|_v$ and a matrix norm $\|\cdot\|_m$ to measure componentwise perturbations to a and $V[a]^{-1}$, we define the structured componentwise condition number with respect to inversion for the primal problem to be

$$\text{(5.1)}\quad \text{invcond}_{v,m}^{\text{primal}}(a) := \lim_{\varepsilon \to 0} \sup_{\|\Delta a\|_v \leqq \varepsilon} \left\{ \frac{\|\widehat{\Delta \text{Inv}}\|_m}{\varepsilon} : \Delta\text{Inv} = V[a+\Delta a]^{-1} - V[a]^{-1} \right\},$$

where $(\widehat{\Delta\text{Inv}})_{ij} = (\Delta\text{Inv})_{ij}/\gamma_{ij}$, and the γ_{ij} are nonnegative tolerances.

Using (2.2), a standard result (see, for example, [7, Lemma 2.3.3]) shows that

$$V[a+\Delta a]^{-1}-V[a]^{-1}=-V[a]^{-1}\,V'[a]\,\mathrm{diag}(\Delta a)\,V[a]^{-1}+O(\varepsilon^2). \tag{5.2}$$

Now, with $\|\widehat{\Delta a}\|_\infty \leqq \varepsilon$,

$$\begin{aligned} |V[a]^{-1}\,V'[a]\,\mathrm{diag}(\Delta a)\,V[a]^{-1}| &= |V[a]^{-1}\,V'[a]\,D_\alpha\,\mathrm{diag}(\widehat{\Delta a})\,V[a]^{-1}| \\ &\leqq |V[a]^{-1}\,V'[a]|\,D_\alpha |V[a]^{-1}|\,\varepsilon, \end{aligned} \tag{5.3}$$

giving a componentwise bound on the perturbation to the inverse. We will define a matrix norm $\|\cdot\|_m$ to be *monotone* if $|A|\leqq|B| \Rightarrow \|A\|_m \leqq \|B\|_m$[1]. It then follows from (5.2) and (5.3) that when a monotone matrix norm is used

$$\mathrm{invcond}^{\mathrm{primal}}_{\infty,m}(a) \leqq \|\widehat{M[a]}\|_m,$$

where $M[a] := |V[a]^{-1}\,V'[a]|\,D_\alpha |V[a]^{-1}|$.

It is possible to get an exact expression when the norm is

$$\|A\|_{\max} := \max_{ij} |a_{ij}|.$$

We make use of the following lemma.

Lemma 5.1. *For any* $A, B \in \mathbb{R}^{N\times N}$

$$\max_{\|\widehat{\Delta a}\|_\infty \leqq \varepsilon} \|A\,\mathrm{diag}(\Delta a)\,B\|_{\max} = \|\,|A|\,D_\alpha |B|\,\|_{\max}\,\varepsilon, \tag{5.4}$$

where $D_\alpha = \mathrm{diag}(\alpha) \geqq 0$ *and* $\Delta a = D_\alpha \widehat{\Delta a}$.

Proof. We have

$$|(A\,\mathrm{diag}(\Delta a)\,B)_{ij}| = \left|\sum_{k=1}^{n} a_{ik}\,\widehat{\Delta a_k}\,b_{kj}\right| \leqq \varepsilon \sum_{k=1}^{n} |a_{ik}|\,\alpha_k\,|b_{kj}| = \varepsilon(|A|\,D_\alpha |B|)_{ij}, \tag{5.5}$$

giving "$\leqq$" in (5.4). If the (r,s) element of $|A|\,D_\alpha |B|$ has maximum modulus then taking $\Delta a_i = \pm \alpha_i \varepsilon$, it is possible to choose the signs so that we have equality in (5.5) for $i=r, j=s$. □

Applying the lemma to (5.2) shows that

$$\mathrm{invcond}^{\mathrm{primal}}_{\infty,\max}(a) = \|\widehat{M[a]}\|_{\max}.$$

The analogous condition number for the dual problem

$$\mathrm{invcond}^{\mathrm{dual}}_{v,m}(a) := \lim_{\varepsilon\to 0} \sup_{\|\widehat{\Delta a}\|_v \leqq \varepsilon} \left\{ \frac{\|\widehat{\Delta \mathrm{Inv}}\|_m}{\varepsilon} : \Delta\mathrm{Inv} = V[a+\Delta a]^{-T} - V[a]^{-T} \right\}, \tag{5.6}$$

---

[1] Monotone vector norms are defined, for example, in [15] and by analogy with Theorem 5.5.10 of [15] it can be shown that a matrix norm is monotone if and only if it is *absolute*; that is, $\|A\| = \|\,|A|\,\|$.

can be analysed in the same manner, and the matrix $M[a]^{\mathrm{T}}$ plays the role of $M[a]$.

The results are collected in the next theorem.

Theorem 5.1. *In the notation above we have*

$$\operatorname{invcond}_{\infty,\max}^{\text{primal}}(a)=\|\widehat{M[a]}\|_{\max},\qquad \operatorname{invcond}_{\infty,\max}^{\text{dual}}=\|\widehat{M[a]}^{\mathrm{T}}\|_{\max},$$

and for any monotone matrix norm $\|\cdot\|_m$,

$$\operatorname{invcond}_{\infty,m}^{\text{primal}}(a)\leqq\|\widehat{M[a]}\|_m,\qquad \operatorname{invcond}_{\infty,m}^{\text{dual}}(a)\leqq\|\widehat{M[a]}^{\mathrm{T}}\|_m.$$

6 Least squares

In this section we look at condition numbers and backward errors for the least squares analogues of (1.2) and (1.3). Here the dual problem naturally arises when low degree polynomials are fitted to discrete data, and hence we examine the condition of the dual case first. We thus assume that functions $\{p_{ij}\}_{i=0,j=0}^{n,m}$ and points $a_0\leqq a_1\leqq\ldots\leqq a_m$ are given, with $m>n$, so that $V[a]\in\mathbb{R}^{(n+1)\times(m+1)}$ in (1.1). We suppose that $V[a]$ has full rank, and that each p_{ij} has a continuous second derivative around a_j. The associated dual least squares problem is

$$\min_{x\in\mathbb{R}^{n+1}}\|V[a]^{\mathrm{T}}x-b\|_2. \tag{6.1}$$

The solution x satisfies the normal equations

$$V[a]\,V[a]^{\mathrm{T}}x=V[a]\,b. \tag{6.2}$$

The corresponding componentwise structured condition number may be defined as

$$\begin{aligned}\operatorname{LScond}_v^{\text{dual}}(a,b):=\lim_{\varepsilon\to 0}\ &\sup_{\|\widehat{\Delta a}\|_v\leqq\varepsilon,\ \|\widehat{\Delta b}\|_v\leqq\varepsilon}\\ &\cdot\left\{\frac{\|\widehat{\Delta x}\|_v}{\varepsilon}:\ V[a+\Delta a]\,V[a+\Delta a]^{\mathrm{T}}(x+\Delta x)\right.\\ &\left.=V[a+\Delta a]\,(b+\Delta b)\right\}.\end{aligned}$$

This quantity can be analysed in a similar way to the condition numbers in Sect. 2. Using the linearisation (2.2) in the constraint

$$V[a+\Delta a]\,V[a+\Delta a]^{\mathrm{T}}(x+\Delta x)=V[a+\Delta a]\,(b+\Delta b)$$

leads to the equation

$$V[a]\,V[a]^{\mathrm{T}}\Delta x=V[a]\,\Delta b+V'[a]\,\operatorname{diag}(\Delta a)\,r_{\mathrm{LS}}-V[a]\,\operatorname{diag}(\Delta a)\,z+O(\varepsilon^2). \tag{6.3}$$

Here $z := V'[a]^T x$ (as in Sect. 2) and r_{LS} denotes the least squares residual, $r_{LS} := b - V[a]^T x$. We may re-write (6.3) as

$$V[a]\,V[a]^T \Delta x = V[a]\,\Delta b + \tilde{V}[a, b, x]\,\Delta a + O(\varepsilon^2),$$

where $\tilde{V}[a, b, x] := V'[a]\,\mathrm{diag}(r_{LS}) - V[a]\,\mathrm{diag}(z)$, and so, in terms of the scaled perturbations,

$$\widehat{\Delta x} = D_\xi^{-1}(V[a]\,V[a]^T)^{-1}(V[a]\,D_\beta \widehat{\Delta b} + \tilde{V}[a, b, x]\,D_\alpha \widehat{\Delta a}) + O(\varepsilon^2).$$

This is a similar expression to (2.5), and by taking norms for a general $\|\cdot\|_v$ and using Lemma 2.1 for $\|\cdot\|_\infty$, we obtain the results in the next theorem.

Theorem 6.1. *In the notation above and for any vector norm, the structured componentwise condition number for the least squares dual problem satisfies*

$$\text{(6.4)}\qquad \mathrm{LScond}_v^{\mathrm{dual}}(a, b) \leqq \|D_\xi^{-1}\,V[a]^{T^+}\,D_\beta\|_v + \|D_\xi^{-1}(V[a]\,V[a]^T)^{-1}\,\tilde{V}[a, b, x]\,D_\alpha\|_v,$$

where $V[a]^{T^+} := (V[a]\,V[a]^T)^{-1}\,V[a]$ *(the pseudo-inverse of* $V[a]^T$*), and this upper bound is not more than twice* $\mathrm{LScond}_v^{\mathrm{dual}}(a, b)$*. For the infinity vector norm,*

$$\text{(6.5)}\qquad \mathrm{LScond}_\infty^{\mathrm{dual}}(a, b) = \||D_\xi^{-1}\,V[a]^{T^+}\,D_\beta| + |D_\xi^{-1}(V[a]\,V[a]^T)^{-1}\,\tilde{V}[a, b, x]\,D_\alpha|\|_\infty.$$

In computing the expressions in (6.4) and (6.5), forming the relevant matrices in a straightforward manner costs $O(n^2 m + mn^2)$ operations, since $V[a]\,V[a]^T$ must be computed and factorised, and $O(m)$ linear systems must then be solved. However, for the one norm or the infinity norm a much cheaper estimate is available using the algorithm in [13]. For example, if the original least squares problem has already been solved by a factorisation method (such as QR) then each matrix-vector product required by [13] costs only $O(mn)$ operations.

Componentwise analysis for the unstructured least squares problem can be found in [3] and [14]. It is interesting to note that whilst there does not appear to exist an exact expression for the componentwise condition number of the unstructured problem for any norm, the $\|\cdot\|_\infty$ dual condition number does have a neat characterisation.

For completeness, we also derive the corresponding result for the least squares primal condition number

$$\mathrm{LScond}_v^{\mathrm{primal}}(a, b) := \lim_{\varepsilon \to 0} \sup_{\|\widehat{\Delta a}\|_v \leqq \varepsilon,\ \|\widehat{\Delta b}\|_v \leqq \varepsilon} \cdot \left\{ \frac{\|\widehat{\Delta x}\|_v}{\varepsilon} : V[a+\Delta a]^T\,V[a+\Delta a]\,(x+\Delta x) = V[a+\Delta a]^T\,(b+\Delta b) \right\},$$

where $V[a]$ in (1.1) is now assumed to be in $\mathbb{R}^{(m+1)\times(n+1)}$, with $m > n$. Linearising the constraint, and writing the result in terms of the scaled perturbations, gives

$$\widehat{\Delta x} = D_\xi^{-1}(V[a]^T\,V[a])^{-1}(V[a]^T\,D_\beta \widehat{\Delta b} + \tilde{\tilde{V}}[a, b, x]\,D_\alpha \widehat{\Delta a}) + O(\varepsilon^2),$$

where

$$\tilde{\tilde{V}}[a,b,x] := \mathrm{diag}(V'[a]^{\mathrm{T}}(b - V[a]\,x)) - V[a]^{\mathrm{T}}\,V'[a]\,\mathrm{diag}(x).$$

Hence, we have

Theorem 6.2. *In the notation above and for any vector norm, the structured componentwise condition number for the least squares primal problem satisfies*

$$\text{(6.6)}\qquad \mathrm{LScond}_v^{\mathrm{primal}}(a,b) \leq \|D_\xi^{-1}\,V[a]^+\,D_\beta\|_v + \|D_\xi^{-1}(V[a]^{\mathrm{T}}\,V[a])^{-1}\,\tilde{\tilde{V}}[a,b,x]\,D_\alpha\|_v,$$

where $V[a]^+ := (V[a]^{\mathrm{T}}\,V[a])^{-1}\,V[a]^{\mathrm{T}}$ *(the pseudo-inverse of* $V[a]$*), and this upper bound is not more than twice* $\mathrm{LScond}_v^{\mathrm{primal}}(a,b)$*. For the infinity vector norm,*

$$\text{(6.7)}\qquad \mathrm{LScond}_\infty^{\mathrm{primal}}(a,b) = \||D_\xi^{-1}\,V[a]^+\,D_\beta| + |D_\xi^{-1}(V[a]^{\mathrm{T}}\,V[a])^{-1}\,\tilde{\tilde{V}}[a,b,x]\,D_\alpha|\|_\infty.$$

It is also possible to analyse the componentwise structured backward error in a least squares solution. Suppose $\tilde{x}$ is an approximate solution to (6.1). We seek perturbations Δa and Δb such that

$$V[a+\Delta a]\,V[a+\Delta a]^{\mathrm{T}}\,\tilde{x} = V[a+\Delta a]\,(b+\Delta b).$$

Linearising this constraint, using (2.2), and re-writing in terms of the scaled perturbations, gives

$$[(V[a]\,\mathrm{diag}(\tilde{z}) - V'[a]\,\mathrm{diag}(\tilde{r}_{\mathrm{LS}}))\,D_\alpha,\ -V[a]\,D_\beta]\begin{bmatrix}\widehat{\Delta a}\\ \widehat{\Delta b}\end{bmatrix} = V[a]\,\tilde{r}_{\mathrm{LS}},$$

where $\tilde{r}_{\mathrm{LS}} := b - V[a]^{\mathrm{T}}\,\tilde{x}$, the residual in $\tilde{x}$. This underdetermined system of $n+1$ equations in $2(m+1)$ unknowns has the same form as (3.11), and the comments there concerning the computation of the exact $\|\cdot\|_\infty$ linearised backward error and upper bounds for the $\|\cdot\|_2$ and $\|\cdot\|_1$ linearised backward errors also apply here.

Similarly, if $\tilde{x}$ is an approximate solution to the primal least squares problem then the linearised constraints for the backward error can be written

$$[(V[a]^{\mathrm{T}}\,V'[a]\,\mathrm{diag}(\tilde{x}) - \mathrm{diag}(V'[a]^{\mathrm{T}}\,\tilde{r}_{\mathrm{LS}}))\,D_\alpha,\ -V[a]^{\mathrm{T}}\,D_\beta]\begin{bmatrix}\widehat{\Delta a}\\ \widehat{\Delta b}\end{bmatrix} = V[a]^{\mathrm{T}}\,\tilde{r}_{\mathrm{LS}},$$

where $\tilde{r}_{\mathrm{LS}}$ now denotes the primal residual, $b - V[a]\,\tilde{x}$.

Acknowledgements. DJH and SGB acknowledge the support of the University of Dundee Research Initiatives Fund and Deutscher Akademischer Austauschdienst, respectively. We thank Nick Higham for valuable comments on this manuscript.

References

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

2. Bartels, S.G. (1991): Two topics in matrix analysis: structured sensitivity for Vandermonde-like systems and a subgradient method for matrix norm estimation. M.Sc. Thesis, University of Dundee
3. Björck, Å. (1991): Component-wise perturbation analysis and error bounds for linear least squares solutions. BIT **31**, 238–244
4. Björck, Å, Pereyra, V. (1970): Solution of Vandermonde systems of equations. Math. Comput. **24**, 893–903
5. Bunch, J.R., Demmel, J.W., Van Loan, C.F. (1989): The strong stability of algorithms for solving symmetric linear systems. SIAM J. Matrix Anal. Appl. **10**, 494–499
6. Gohberg, I., Koltracht, I. (1990): On the inversion of Cauchy matrices. In: M.A. Kaashoek, J.H. van Schuppen, A.C.M. Ran, eds., Proceedings of Signal Processing, Scattering and Operator Theory, and Numerical Methods. Birkhäuser, Basel, pp. 381–392
7. Golub, G.H., Van Loan, C.F. (1989): Matrix Computations, 2nd ed. Johns Hopkins University Press, Baltimore, Md.
8. Hager, W.W. (1984): Condition estimates. SIAM J. Sci. Statist. Comput. **5**, 311–316
9. Higham, D.J., Higham, N.J. (1992): Backward error and condition of structured linear systems. SIAM J. Matrix Anal. Appl. **13**, 162–175
10. Higham, D.J., Higham, N.J. (1992): Componentwise perturbation theory for linear systems with multiple right-hand sides. Linear Algebra Appl. (to appear)
11. Higham, N.J. (1987): Error analysis of the Björck-Pereyra algorithms for solving Vandermonde systems. Numer. Math. **50**, 613–632
12. Higham, N.J. (1990): Stability analysis of algorithms for solving confluent Vandermonde-like systems. SIAM J. Matrix Anal. Appl. **11**, 23–41
13. Higham, N.J. (1988): FORTRAN codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation (Algorithm 674). ACM Trans. Math. Soft. **14**, 381–396
14. Higham, N.J. (1990): Computing error bounds for regression problems. In: P.J. Brown, W.A. Fuller, eds., Statistical Analysis of Measurement Error Models and Applications. Contemporary Mathematics 112. Amer. Math. Soc., pp. 195–208
15. Horn, R.A., Johnson, C.R. (1985): Matrix Analysis. Cambridge University Press, Cambridge
16. Oettli, W., Prager, W. (1964): Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. Numer. Math. **6**, 405–409
17. Rohn, J. (1989): New condition numbers for matrices and linear systems. Computing **41**, 167–169
18. Skeel, R.D. (1979): Scaling for numerical stability in Gaussian elimination. J. Assoc. Comput. Mach. **26**, 494–526
19. Pham Dinh Tao (1984): Convergence of a subgradient method for computing the bound norm of matrices. Linear Algebra Appl. **62**, 163–182 [in French]
20. Traub, J.F. (1966): Associated polynomials and uniform methods for the solution of linear problems. SIAM Review **8**, 277–301