Applied Mathematics E-Notes, 7(2007), 65-70 © ISSN 1607-2510
Available free at mirror sites of http://www.math.nthu.edu.tw/∼amen/

On The Belonging Of A Perturbed Vector To A Subspace From A Numerical View Point*

Claudia Fassino†

Received 17 June 2005

Abstract

We propose a criterion in order to establish, dealing with perturbed data, when a vector v belongs to a subspace W of a vector space U from a numerical point of view. The criterion formalizes the intuitive idea that, due to the errors that affect the knowledge of the vector and the subspace, we consider the exact vector lying in the exact subspace if a "sufficiently small perturbation" of v belongs to a "sufficiently small perturbation" of W. The criterion, obviously chosen as independent of the norm of v and of the basis of W, is computationally simple to use, but it requires the choice of a threshold. Suitable values of the threshold such that the criterion adheres to the previous intuitive idea are found both when we know a basis of the exact subspace and when an orthonormal basis of the perturbed subspace is given.

1 Introduction

Several mathematical problems involve the necessity of detecting if an assigned vector v_0 of a vector space U lies in an assigned subspace W_0 of U. When v_0 and W_0 are known in an exact way, the theoretical approach to this detection gives a well known completely satisfactory answer.

However, when we deal with "real" data, for example with physical measurements, the coordinates of the vectors are, in general, affected by errors, whose estimation is known. Obviously, in this case, the exact approach can lead to meaningless answers, especially when the vector is close to lying in the subspace.

In this paper we present a suitable criterion for detecting "numerical" dependence. This criterion is powerful enough to overcome the difficulties occurring in presence of perturbed data, in the sense that if the perturbed vector v does not belong from the numerical point of view to the perturbed subspace W, we show that also the exact vector v_0 does not lie in the exact subspace W_0. Vice versa, if v "numerically" lies in W, we prove that there exists at least one small perturbation $\hat{v}$ of v exactly belonging to the subspace $\widehat{W}$, which is a slight perturbation of W. In this case, since $\hat{v}$ and $\widehat{W}$ could correspond to the exact vector v_0 and the exact subspace W_0, obviously unknown, the

*Mathematics Subject Classifications: 15A03, 15A18, 65F99, 65G99.

†Dipartimento di Matematica, Università di Genova, via Dodecaneso 35, 16146 Genova, Italy.

criterion points out the possibility for v_0 of belonging exactly to W_0. Therefore the criterion formalizes the intuitive idea that we consider the exact vector v_0 lying in W_0 from a numerical point of view if v is "sufficiently near" to W and vice versa.

The criterion, presented in Section 2, has been obviously chosen as independent of the norm of v and of the basis of the subspace W; moreover, it is computationally simple to use. However, the criterion requires the choice of a suitable threshold in order to adhere to the intuitive idea described above. We also show, in Section 3, how to choose the threshold both in the simpler case, when a basis of the exact subspace W_0 is known and when a set of orthonormal perturbed vectors, which is regarded as the basis of the perturbed subspace W, is given.

We specify that our interest in this problem arises from the analysis of the Buchberger-Möller algorithm [1] for computing the Gröbner basis of an ideal of polynomials of $\mathbf{Q}[x_1, \ldots, x_k]$, vanishing on a finite set of perturbed points of $\mathbf{Q}^k$. At each step, the numerical version of the Buchberger-Möller algorithm executes a fundamental check based on the linear dependence of a perturbed vector v on a set of orthonormal perturbed vectors $v_1, \ldots, v_n$, and so we need a criterion for the "numerical" belonging of a vector to a subspace. Then, since this algorithm is implemented using the CoCoA package [2], which performs all the computations on the rational numbers in exact arithmetic, we do not deal with roundoff errors and we are only interested in the effects of the errors which perturb the input points. Moreover, for the same reason, the vectors $v_1, \ldots, v_n$, computed from perturbed data, are exactly orthonormal and so we are interested in the case where an orthonormal basis of the perturbed subspace is known.

2 On the "Numerical" Dependence

Intuitively, when the exact case is unknown, we suppose the exact vector v_0 lying in the exact subspace W_0 if a perturbation v of v_0 is "near" to a perturbation W of W_0. We formalize this idea analyzing the ratio $\|\rho\|/\|v\|$, where ρ is the component of v orthogonal to W and $\|\cdot\|$ is the vector 2-norm.

DEFINITION 2.1. Given a subspace W and a threshold $S > 0$, a vector v "numerically" lies in W if

$$\frac{\|\rho\|}{\|v\|} \leq S, \tag{1}$$

where ρ is the component of v orthogonal to W. In this case we write $v \in_S W$.

Condition (1) obviously formalizes the intuitive idea described above in a way dependent only on the subspace W and on the direction of v and then independently of $\|v\|$ and of the choice of the basis of W. Therefore, we can suppose $\|v\| = 1$ and W described by an orthonormal basis, if necessary. Moreover, note that, given a basis $\{v_1, \ldots, v_n\}$ of W, the vector ρ can be obtained computing the residual of a least squares problem with matrix $[v_1, \ldots, v_n]$ and right-side vector v. However, condition (1) depends on the choice of the threshold S. Of course, an arbitrary choice of S can lead to results that do not reflect the intuitive idea. Then it is necessary to analyze how the threshold can be chosen in relation with the errors on the data. In the next section, we show how to choose a suitable value of S, in both situations, when we know the exact subspace W_0 and when we know only an approximation of it.

3 The Choice of Suitable Thresholds

We analyze firstly the simpler case when the exact subspace $W_0 \subset U$ is known. Since the vector v is perturbed by errors, it can be considered as a representative element of a set I which contains the unknown exact vector v_0 and its "admissible perturbations". Given an estimation ε of the relative error which perturbs v componentwise, we denote by I the set of "admissible perturbations" of v

$$I = \{t \in U \ : \ \|t - v\| \leq \varepsilon \|v\|\}.$$

Following the intuitive idea of "numerical" dependence, a suitable threshold S must be such that if $v \notin_S W_0$ then each vector $t \in I$ does not lie in W_0, that implies $v_0 \notin W_0$. Vice versa, if $v \in_S W_0$ there exists a vector $\hat{t}$ of I which belongs to W_0 so that it can happen that v_0 lies in W_0.

We want to show that, if W_0 is known, the value $S = \varepsilon$ is a suitable choice. If $\{v_1^{(0)}, \ldots, v_n^{(0)}\}$ is a basis of the exact subspace W_0, then v can be written as $v = \sum_{i=1}^n \alpha_i v_i^{(0)} + \rho$, where $\rho \in W_0^{\perp}$.

Let $v \notin_\varepsilon W_0$: then, for condition (1), $\|\rho\| > \varepsilon\|v\|$. Since $\rho \in W_0^{\perp}$, for each $t \in W_0$, $t = \sum_{i=1}^n \beta_i v_i^{(0)}$, we have

$$\|v - t\|^2 = \| \sum_{i=1}^n (\alpha_i - \beta_i) v_i^{(0)} - \rho\|^2 = \| \sum_{i=1}^n (\alpha_i - \beta_i) v_i^{(0)} \|^2 + \|\rho\|^2.$$

We have then $\|v - t\| \geq \|\rho\| > \varepsilon\|v\|$, and therefore $t \notin I$.

Vice versa, if $v \in_\varepsilon W_0$, that is $\|\rho\| \leq \varepsilon\|v\|$, the vector $\hat{t} = v - \rho$, belonging to W_0, is such that $\|v - \hat{t}\| = \|\rho\| \leq \varepsilon\|v\|$. It follows that there exists an element of I, the vector $\hat{t}$, which also lies in W_0.

The more general situation where not only the exact vector v_0, but also the exact subspace W_0 is unknown, requires a different analysis. We suppose to know a set of orthonormal vectors $\{v_1, \ldots, v_n\}$, approximation of the orthonormal basis $\{v^{(0)}, \ldots, v_n^{(0)}\}$ of W_0, and a perturbed vector v, with $\|v\| = 1$, approximation of v_0. Denoting by W the subspace spanned by $\{v_1, \ldots, v_n\}$, we apply condition (1) for detecting if the perturbed vector v "numerically" belongs to the perturbed subspace W and, as above, we have to choose a suitable threshold S. With respect to the previous case, there is a conceptual difference because in this situation we have to analyze also the effects of small perturbations of the basis of W. For this reason, a suitable threshold is such that if $v \notin_S W$, the vectors $\{v_1, \ldots, v_n, v\}$ are linearly independent even if all of them are slightly perturbed. Vice versa, if $v \in_S W$, we require that there exists a small perturbation of $\{v_1, \ldots, v_n, v\}$ which is a set of linear dependent vectors.

We suppose that ε is an estimation of the componentwise relative errors such that $\|V - V_0\| \leq \varepsilon\|V\|$, where $V_0 = [v_1^{(0)}, \ldots, v_n^{(0)}, v_0]$ and $V = [v_1, \ldots, v_n, v]$ are, respectively, the matrices formed by the exact and the perturbed vectors, and $\| \ \|$ is the 2-norm of a matrix. We define the set of the “admissible perturbations” as:

$$\mathcal{I} = \{B \mid \|V - B\| \leq \varepsilon \|V\|\}.$$

A suitable threshold must be such that, if $v \notin_S W$, all the elements of $\mathcal{I}$ are full rank matrices, and so, since $V_0 \in \mathcal{I}$, the exact vectors $v_1^{(0)}, \ldots, v_n^{(0)}, v_0$ are linearly independent. Vice versa, if $v \in_S W$, the value of S must be such that a deficient rank matrix $\hat{B} \in \mathcal{I}$ exists. Since $\hat{B}$ can correspond to V_0, the exact vectors $v_1^{(0)} \ldots v_n^{(0)}$ and v_0 could be linearly dependent.

It is well known [3] that, if σ_{n+1} is the least singular value of V, then

$$\sigma_{n+1} = \min_{rank(B)=n} \|B - V\|.$$

Therefore a suitable threshold must be such that the condition $v \notin_S W$ implies

$$\varepsilon \|V\| < \sigma_{n+1}, \tag{2}$$

ensuring that all the elements of $\mathcal{I}$ are full rank matrices. Vice versa, the condition $v \in_S W$ must imply $\varepsilon \|V\| \geq \sigma_{n+1}$. Since the largest singular value σ_1 of V is $\sigma_1 = \|V\|$, the inequality (2) can be rewritten as $\varepsilon \sigma_1 < \sigma_{n+1}$. Condition (2) points out that, when the relative error ε is greater than 1, the problem of belonging of v to W is intrinsically ill posed. In fact in this case $\varepsilon \|V\| = \varepsilon \sigma_1 \geq \sigma_{n+1}$ and then, independently of the choice of S, always there exists a rank deficient matrix $\widehat{B} \in \mathcal{I}$. Therefore later on we suppose $\varepsilon < 1$.

The following theorem gives the singular values of the matrix V.

THEOREM 3.1. Let $\{v_1, \ldots, v_n\}$ be a set of orthonormal vectors in $\mathbf{R}^m$, $m > n$, and let $v \in \mathbf{R}^m$ be a vector such that

$$v = \sum_{i=1}^{n} \alpha_i v_i + \rho \ \text{ with } \ \rho \in Span\{v_1, \ldots, v_n\}^{\perp}.$$

Then the singular values of the matrix $V = [v_1, \ldots, v_n, v]$ are

$$\begin{aligned}
\sigma_1 &= \frac{1}{\sqrt{2}} \sqrt{1 + \|v\|^2 + \sqrt{(1 + \|v\|^2)^2 - 4\|\rho\|^2}} \geq 1, \\
\sigma_2 &= \cdots = \sigma_n = 1, \\
\sigma_{n+1} &= \frac{1}{\sqrt{2}} \sqrt{1 + \|v\|^2 - \sqrt{(1 + \|v\|^2)^2 - 4\|\rho\|^2}} \leq 1.
\end{aligned}$$

PROOF. It is well known [4] that the values $\sigma_1^2, \ldots, \sigma_{n+1}^2$ are the eigenvalues of the matrix $M = V^t V$,

$$M = \begin{pmatrix} I_n & \alpha \\ \alpha^t & \|v\|^2 \end{pmatrix},$$

where $\alpha^t = (\alpha_1 \ldots \alpha_n)$ and I_n is the $n \times n$ identity matrix.

Let M_j be the principal submatrix of M formed by selecting the last j rows and the last j columns, and let d_j be the determinant of $(M_j - \lambda I_j)$. If we compute d_{n+1} with respect to the first row, we have

$$d_{n+1} = (1-\lambda)d_n - \alpha_1^2(1-\lambda)^{n-1}$$

and, in an analogous way,

$$d_{j+1} = (1-\lambda)d_j - \alpha_{n-j+1}^2(1-\lambda)^{j-1},$$

so that

$$d_{n+1} = (1-\lambda)^n d_1 - \|\alpha\|^2(1-\lambda)^{n-1}.$$

Since $d_1 = \|v\|^2 - \lambda$ and $\|v\|^2 = \|\rho\|^2 + \|\alpha\|^2$, we obtain

$$d_{n+1} = (1-\lambda)^{n-1}\left[\lambda^2 - (1+\|v\|^2)\lambda + \|\rho\|^2\right],$$

so that the eigenvalues of V^tV are 1, with algebraic multiplicity $n-1$, and

$$\begin{aligned} \lambda_+ &= \frac{1}{2}\left[1 + \|v\|^2 + \sqrt{(1+\|v\|^2)^2 - 4\|\rho\|^2}\right], \\ \lambda_- &= \frac{1}{2}\left[1 + \|v\|^2 - \sqrt{(1+\|v\|^2)^2 - 4\|\rho\|^2}\right]. \end{aligned}$$

Since $\|\rho\|^2 \le \|v\|^2$, it is easy to verify that $\lambda_+ \ge 1$ and $\lambda_- \le 1$, from which the theorem follows.

As consequence of Theorem 3.1, since $\|v\| = 1$, the inequality (2) can be rewritten as

$$\varepsilon\sqrt{1 + \sqrt{1 - \|\rho\|^2}} < \sqrt{1 - \sqrt{1 - \|\rho\|^2}},$$

and therefore, when $\varepsilon < 1$, it is equivalent to

$$\|\rho\| > \frac{2\varepsilon}{1+\varepsilon^2}.$$

The previous analysis indicates that the value $S = 2\varepsilon/(1+\varepsilon^2)$ is a suitable choice for the threshold in this more general case.

4 Conclusions

We have proposed a threshold dependent criterion for establishing when a vector v belongs to a subspace W from a numerical point of view if v and, in a more general case, W are perturbed by errors. When an estimation ε of the relative error on the data is known, we show that the following values of S are suitable threshold:

i) $S = \varepsilon$ if a basis of the exact subspace is known,

ii) $S = 2\varepsilon/(1+\varepsilon^2)$ if an orthonormal basis of the perturbed subspace is given and $\|v\| = 1$.

Several numerical tests straightforwardly confirm, from the computationally point of view, the theoretical evaluations (i) and (ii) of the threshold S.

References

[1] B. Buchberger and H. Möller, The construction of multivariate polynomials with preassigned zeros, in: Proceedings of the European Computer Algebra Conference (EUROCAM 82), Vol. 144 of Lectures Notes In Comp. Sci., Springer, Marseille - France, 1982, pp. 24–31.

[2] J. Abbott, A. Bigatti, M. Kreuzer and L. Robbiano, Computing ideals of points, JSYMC 30(4)(2000) 341–356.

[3] Å. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.

[4] G. H. Golub and C. V. Loan, Matrix Computations. Second Edition, The Johns Hopkins University Press, Baltimore, Maryland, 1991.