

ON GENERALIZED HÖLDER INEQUALITY

WEI H. YANG

Department of Mechanical Engineering and Applied Mechanics, University of Michigan, Ann Arbor,
MI 48109, U.S.A.

(Received 15 February 1990; received for publication 30 July 1990)

Key words and phrases: Inequalities, vector and matrix norms, functional analysis, plasticity.

INTRODUCTION

A FAMILY of inequalities concerning inner products of vectors and functions began with Cauchy. The extensions and generalizations later led to the inequalities of Schwarz, Minkowski and Hölder. The well known Hölder inequality involves the inner product of vectors measured by Minkowski norms. In this paper, another step of extension is taken so that a Hölder type inequality may apply to general, paired non-Euclidean norms. We restrict the discussion to finite dimensional spaces. A theorem is established for such generalized Hölder inequality together with its sharpness condition. The new sharpness condition involves the gradient of the norm function. The previous inequalities of this type and their associated sharpness conditions become special cases of this new inequality which is a fundamental primal–dual relation of two spaces defined by paired norms on their respective elements. Given an arbitrarily defined primal norm in one space, a method is presented for constructing the dual norm in the other. The generalized Hölder inequality is the basis for many known minimax (duality) theorems in applied fields. Particular examples of this duality are found in mathematical models of plastic behavior.

Inequalities appear frequently in algebra, geometry and analysis. This form of mathematical statements is so prevalent that it is difficult to state their relevance and applications exhaustively. Inequalities are used to demarcate numbers, vectors, matrices and functions, e.g. to be positive definite; to define sets, norm measures, convexity of sets and functions; to compare and bound functions with another function just to name a few. There are books devoted exclusively to inequalities [1].

A class of inequalities concerning inner products of vectors and functions can be grouped into two frequently encountered ones in literature although one is a special case of the other. The Schwarz inequality applies to the Euclidean and Hilbert spaces [2]. The Hölder inequality extends the concept to certain non-Euclidean and non-Hilbert spaces associated with a specific family of norm measures defined by Minkowski [3]. We shall restrict our discussion to finite dimensional spaces. A further extension of the Hölder inequality is presented in this paper for general paired non-Euclidean spaces.

The familiar Schwarz inequality for the Euclidean space R^n can be stated for the inner product

$$|\mathbf{x}'\mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2, \quad (1)$$

where \mathbf{x}, \mathbf{y} are vectors in R^n , t transposes a vector and $\|\cdot\|_2$ denotes the Euclidean norm. The

equality holds if

$$\mathbf{y} = \alpha \mathbf{x}, \quad \alpha \in R, \tag{2}$$

i.e. the two vectors are colinear. The inequality (1) is often used in the upper bounding process of a mathematical analysis. A ‘‘sharp’’ upper bound that includes the equality case is vitally important in the field of functional analysis [4] and its applications. Obviously, a function that is bounded above has a finite supremum. The maximum of a function is contained in the range of a sharp upper bound function. Therefore a search for the least upper bound (supremum) will help recover the maximum of the original function. This indirect method of finding the maximum of a function will fail if its upper bound function is not sharp. The sharpness condition (2) for the Schwarz inequality does not hold in a general real vector space R^n when non-Euclidean norm measures are used. The co-linearity condition in (2) will have to be modified for it to retain general validity.

In the present paper, we intend to study norms $\|\mathbf{x}\|_{(p)}$, $\|\mathbf{y}\|_{(d)}$ in R^n such that

$$|\mathbf{x}'\mathbf{y}| \leq \|\mathbf{x}\|_{(p)}\|\mathbf{y}\|_{(d)}$$

for all $\mathbf{x}, \mathbf{y} \in R^n$, and to determine conditions under which equality holds in this relation (sharpness). This pair of norms are thought of as having a dual relation implied by their subscript, (p) for primal, and (d) for dual. One of these norms can be arbitrarily defined. We use parentheses on the subscripts p and d to emphasize generality and to prevent confusion with the Minkowski norms.

The following simple example illustrates our point. Consider the two norms in R^2 , $\|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|\}$ and $\|\mathbf{y}\|_1 = |y_1| + |y_2|$. It is well known that $|\mathbf{x}'\mathbf{y}| \leq \|\mathbf{x}\|_\infty\|\mathbf{y}\|_1$ for all $\mathbf{x}, \mathbf{y} \in R^2$, and that equality holds under a certain condition which, in general, is not colinearity of \mathbf{x} and \mathbf{y} . This can be seen in geometric terms by considering two unit vectors $\mathbf{x}, \mathbf{y} \in R^2$ in their respective norms

$$\begin{aligned} \|\mathbf{x}\|_\infty &= \max\{|x_1|, |x_2|\} = 1, \\ \|\mathbf{y}\|_1 &= |y_1| + |y_2| = 1. \end{aligned} \tag{3}$$

Clearly, all such vectors lie on two ‘‘unit circles’’, one in the form of a square for \mathbf{x} and the other in the form of a diamond for \mathbf{y} as shown in Fig. 1. Then

$$|\mathbf{x}'\mathbf{y}| \leq \|\mathbf{x}\|_\infty\|\mathbf{y}\|_1 = 1 \tag{4}$$

holds for any \mathbf{x} and \mathbf{y} defined in (3). Here, the vectors \mathbf{x} and \mathbf{y} are called the pair of dual variables.

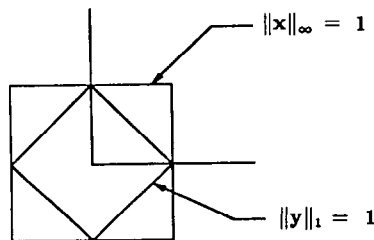


Fig. 1. Unit circles for vectors in R^2 with different norm measures.

Consider a specific vector

$$\mathbf{x} = (1 - \alpha)^t, \quad 0 < \alpha < 1. \tag{5}$$

The vector \mathbf{y} in the direction of this \mathbf{x} has the form

$$\mathbf{y} = \left(\frac{1}{1 + \alpha} \quad \frac{\alpha}{1 + \alpha} \right)^t, \tag{6}$$

so that it has unit length as defined in (3). The inner product of these two colinear vectors can be computed to yield

$$\mathbf{x}^t \mathbf{y} = (1 + \alpha^2)/(1 + \alpha) < 1. \tag{7}$$

Hence, the sharpness condition in (2) is not valid in general. The intuition, that an inner product of two vectors attains maximum (or minimum) when they are colinear, needs qualification. The sharpness condition for the dual vectors defined in (3) does exist and arises from certain specific pairing other than the colinearity of the two vectors. The pairing depends on the norm measures defined. We shall explain such pairing in a slightly more general setting than the 1-norm and ∞ -norm defined in (3).

MINKOWSKI NORMS AND HÖLDER INEQUALITY

The family of Minkowski norms [4] for a vector $\mathbf{x} \in R^n$ is defined by

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p \leq \infty, \tag{8}$$

which, called a p -norm, includes all three norms discussed in the previous section as special cases, $p = 1, 2$ and ∞ . A pair of norms, $\|\cdot\|_p$ and $\|\cdot\|_q$, in this family are said to have a dual relation [4] if

$$\frac{1}{p} + \frac{1}{q} = 1, \quad 1 \leq p, q \leq \infty. \tag{9}$$

The Hölder inequality holds for any pair of such dual norms:

$$|\mathbf{x}^t \mathbf{y}| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q. \tag{10}$$

It is well known that the equality holds for $1 \leq p, q < \infty$ if

$$\text{either } y_i = \alpha |x_i|^{p-1} \text{sign}(x_i) \quad \text{or } x_i = \alpha |y_i|^{q-1} \text{sign}(y_i), \quad \alpha \in R, i = 1, 2, \dots, n.$$

First, let us show that this sharpness result can be simply obtained by requiring that

$$\text{either } \mathbf{y} = \alpha \nabla \|\mathbf{x}\|_p \quad \text{or } \mathbf{x} = \alpha \nabla \|\mathbf{y}\|_q, \tag{11}$$

where ∇ is the gradient operator. When operated on a C^1 function of n variables, ∇ produces a gradient vector in R^n . A gradient vector is normal to the level sets (contours) of the norm function. Thus, the condition for sharpness given in (11) can also be called the normality relation. The vectors satisfying (11) belong to a dual pair. To avoid the issue of lack of differentiability of certain norms, we postpone the discussion on the limiting cases ($p, q = 1, \infty$) in which the norms are C^0 functions.

We shall show the equality case of (10) under one of the normality relations in (11) by first choosing a p -norm ($1 < p < \infty$) for \mathbf{x} then obtaining the components of \mathbf{y} explicitly by

differentiation

$$y_i = k \frac{\partial}{\partial x_i} (\sum |x_j|^p)^{1/p} = k(\sum |x_j|^p)^{1/p-1} |x_i|^{p-1} \text{sign}(x_i), \tag{12}$$

where $k > 0$ is chosen as the proportional factor and the ranges of the sums over j and the index i are implied to span from 1 to n . Using (12), we can form the q -norm of \mathbf{y} from its components and, with some simplification, obtain

$$\|\mathbf{y}\|_q = (\sum |y_i|^q)^{1/q} = k, \tag{13}$$

which gives the proportional factor k an explicit expression. Again using (12) and (13), we can produce the inner product for the pair of dual vectors

$$\mathbf{x}'\mathbf{y} = \sum x_i y_i = \|\mathbf{x}\|_p \|\mathbf{y}\|_q. \tag{14}$$

The equality is thus obtained and the normality relation is the desired sharpness condition. The absolute bound on $\mathbf{x}'\mathbf{y}$ is not needed here since k is chosen to be positive. Here relation (9) was used explicitly in proving relations (13) and (14). It will be demonstrated from the argument during the construction of dual norms in a later section. Inequalities (1) and (4) are special cases of (10).

The normality relation between dual vectors is not restricted to the family of Minkowski norms. This dual structure when applied to general non-Euclidean norms yields an inequality similar to that of Hölder.

GENERALIZED HÖLDER INEQUALITY FOR NON-EUCLIDEAN VECTOR SPACES

All norms share a common property as convex, nonnegative and homogeneous functions of degree one. One may arbitrarily define a norm as long as it satisfies the conditions

- (i) $\|\mathbf{x}\| \geq 0,$
 - (i') $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = 0,$
 - (ii) $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|,$
 - (iii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
- (15)

for all $\alpha \in R$ and $\mathbf{x}, \mathbf{y} \in R^n$. The part (i') is omitted for the definition of a seminorm.

Usually, the definition of a primal norm (or seminorm) arises naturally from a specific application. A dual norm must be matched to the primal norm so that a theorem for the generalized Hölder inequality may be stated below.

THEOREM. For any two vectors $\mathbf{x}, \mathbf{y} \in R^n$ where \mathbf{x} is measured by a properly defined primal norm (or seminorm), $\|\cdot\|_{(p)}$, there exists a dual norm (or seminorm) $\|\cdot\|_{(d)}$ such that the inequality

$$|\mathbf{x}'\mathbf{y}| \leq \|\mathbf{x}\|_{(p)} \|\mathbf{y}\|_{(d)} \tag{16}$$

holds. The case of equality is attained when the sharpness or normality relation

$$\mathbf{y} = \|\mathbf{y}\|_{(d)} \nabla \|\mathbf{x}\|_{(p)} \tag{17}$$

holds.

Here we assume that $\|\mathbf{x}\|_{(p)}$ is a function of class C^1 for $\mathbf{x} \neq 0$. The theorem, of course, covers the original Hölder inequality (10) as a special case. We shall establish the equality case of (16) followed by a proof of the inequality. Then a systematic method to construct the dual norm from a given primal norm will be presented.

Along any ray, $\{\mathbf{x} \in R^n : \mathbf{x} = \alpha \mathbf{c}, \alpha \geq 0\}$ where $\mathbf{c} \in R^n$ is a nonzero constant vector, a norm function being homogeneous of degree one has a constant gradient. Furthermore, $\|\mathbf{x}\| = 0$ when $\mathbf{x} = 0$. Geometrically, a norm function $\|\mathbf{x}\|$ represents a cone in R^{n+1} . All norm functions satisfy

$$\|\mathbf{x}\| = \mathbf{x}'\nabla\|\mathbf{x}\|. \tag{18}$$

Choosing a \mathbf{y} according to (17) and forming the inner product with \mathbf{x} , we obtain after applying (18)

$$\mathbf{x}'\mathbf{y} = \|\mathbf{x}\|_{(p)}\|\mathbf{y}\|_{(d)} \tag{19}$$

which proves the equality case stated in (16). Again, we have chosen a positive inner product. Otherwise, an absolute bound on the inner product will be added to the left-hand side of (19).

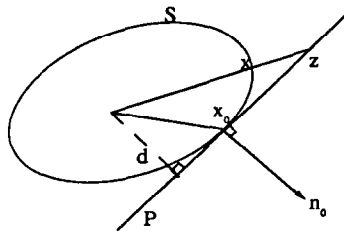


Fig. 2. The unit sphere and its supporting hyperplane.

To prove the general inequality (16), we use the convexity property of the norm functions. Let the “unit sphere” for the primal norm be defined by the set

$$S = \{\mathbf{x} \in R^n : \|\mathbf{x}\|_{(p)} = 1\} \tag{20}$$

as shown in Fig. 2 where $\mathbf{x}_0 \in S$ is an arbitrary vector (or a point).

At the point \mathbf{x}_0 , there is a normal vector

$$\mathbf{n}_0 = \nabla\|\mathbf{x}\|_{(p)}|_{\mathbf{x}=\mathbf{x}_0} \tag{21}$$

and a hyperplane $P \subset R^n$ passing through \mathbf{x}_0 and normal to \mathbf{n}_0 . Then P is the supporting hyperplane of S at \mathbf{x}_0 since the entire set S lies on one side of P .

Since all norms are homogeneous functions of degree one, the origin \mathbf{O} of R^n lies strictly inside of the unit sphere in Fig. 2. Therefore the origin and the set S lie on the same side of P . A line from the origin to a point $\mathbf{z} \in P$ will intersect the unit sphere. Let this intersection be denoted by $\mathbf{x} \in S$ as shown. Again using the properties of convexity and homogeneity of degree one of the primal norm, it is easy to deduce that this intersection is unique and satisfies

$$1 = \|\mathbf{x}_0\|_{(p)} = \|\mathbf{x}\|_{(p)} \leq \|\mathbf{z}\|_{(p)} \quad \text{and} \quad \mathbf{x}'\mathbf{n}_0 \leq \mathbf{z}'\mathbf{n}_0 = \mathbf{x}_0'\mathbf{n}_0 = d, \tag{22}$$

where d is the projection of all vectors in P in the direction of \mathbf{n}_0 .

Let $\mathbf{y}_0 = k\mathbf{n}_0(k > 0)$ be the dual vector that satisfies the normality relation (11) associated with the primal vector \mathbf{x}_0 . Then

$$\mathbf{x}'\mathbf{y}_0 \leq \mathbf{z}'\mathbf{y}_0 = \mathbf{x}'_0\mathbf{y}_0 = \|\mathbf{x}_0\|_{(p)}\|\mathbf{y}_0\|_{(d)} = \|\mathbf{x}\|_{(p)}\|\mathbf{y}_0\|_{(d)} \tag{23}$$

such that only the dual vectors \mathbf{x}_0 and \mathbf{y}_0 which satisfy the normality relation (17) achieve equality. All other dual pairs \mathbf{x} and \mathbf{y} satisfy inequality (16). Since the vector $\mathbf{x}_0 \in S$ is arbitrarily chosen, the theorem is proven.

CONSTRUCTING THE DUAL NORM

The specific form of $\|\mathbf{y}\|_{(d)}$ can be determined from a given $\|\mathbf{x}\|_{(p)}$. If the primal norm is given as a set of numerical data on a “unit sphere” as it may be the case in an engineering problem, then a local gradient can be computed approximately at each data point on the sphere. These gradient vectors will trace a “unit sphere” of the dual norm. This numerical method, although practical, may not be a satisfactory demonstration of the dual norm construction.

We shall assume that the primal norm is a closed-form C^1 function such that its first derivatives can be computed everywhere. Dividing (17) by $\|\mathbf{y}\|_{(d)}$, then the equation represents a map from \mathbf{x} to \mathbf{y} . If $\|\mathbf{x}\|_{(p)}$ is strictly convex, the inverse map exists and the equation (17) can be solved for the components of \mathbf{x} in terms of that of \mathbf{y} . By taking the (p) -norm of \mathbf{x} from its components and setting it to unity, an expression for $\|\mathbf{y}\|_{(d)}$ is obtained in terms of the components of \mathbf{y} . If $\|\mathbf{x}\|_{(p)}$ contains a linear portion, the points on the linear portion maps to a single point on $\|\mathbf{y}\|_{(d)}$. The map is no longer one-to-one but construction of the dual norm can still proceed.

In a previous section, we have verified the Hölder inequality (10) without questioning the origin of the dual relation (9) of the Minkowski family of norms. For a given primal norm $\|\mathbf{x}\|_{(p)} (1 < p < \infty)$, we shall now derive the form of its dual norm $\|\mathbf{y}\|_{(d)}$ without resorting to (9). Rewriting (12) by letting $\|\mathbf{x}\|_p = 1$ and $k = \|\mathbf{y}\|_{(d)}$, we have

$$y_i = \|\mathbf{y}\|_{(d)}|x_i|^{p-1} \text{sign}(x_i) \tag{24}$$

from which we may solve for $|x_i|$

$$|x_i| = \left(\frac{|y_i|}{\|\mathbf{y}\|_{(d)}} \right)^{1/(p-1)} \tag{25}$$

It is an easy matter to form $\|\mathbf{x}\|_p$ from the components x_i and to obtain

$$1 = \|\mathbf{x}\|_p = \|\mathbf{y}\|_{(d)}^{-1/(p-1)} (\sum |y_i|^{p/(p-1)})^{1/p} \tag{26}$$

From the above equation

$$\|\mathbf{y}\|_{(d)} = (\sum |y_i|^{p/(p-1)})^{(p-1)/p} \tag{27}$$

is obtained to give the explicit form of $\|\mathbf{y}\|_{(d)}$ in terms of y_i and the information of the primal norm. Now letting $p/(p - 1) = q$, we confirm the dual relation (9) and conclude

$$\|\mathbf{y}\|_{(d)} = \|\mathbf{y}\|_q, \tag{28}$$

as the dual norm for the primal norm $\|\mathbf{x}\|_p$. Construction of the dual norms for other non-Euclidean C^1 primal norms are analogous.

C^0 NORMS

The cases $p = 1, \infty$ correspond to C^0 norms which are not differentiable in the usual sense at certain values of the independent variable. They can be regarded as the limiting cases and the inequality (4) is also confirmed from the above analysis with the values $q = \infty, 1$ obtained for the corresponding dual norms by taking limits from the p -norms in C^1 . Although there is no mathematical difficulty to demonstrate the generalized Hölder inequality and its normality relation for C^0 norms, it is desirable to offer an interpretation for the theorem at those points where the norm function is not differentiable. We shall call those points the vertices of a norm.

Consider again the example in R^2 shown in Fig. 1. The gradient of $\|x\|_\infty$ has only four distinctive values $(1, 0), (0, 1), (-1, 0)$ and $(0, -1)$ each evaluated along a respective edge DA, AB, BC and CD on the unit circle of the primal norm as shown in Fig. 3a.

These four values are shown as four points A', B', C' and D' in Fig. 3b on the "unit circle" of the dual norm, $\|y\|_{(d)} = 1$, yet to be constructed. Since the dual norm must be a convex and continuous function, these four points must be connected. Let us connect the four points by straight lines to form a diamond $A'B'C'D'$. Then no points of the "unit circle" should fall inside the diamond or the dual norm function would be non-convex. There exist, of course, convex functions passing through these four points but outside the diamond. We shall show that they cannot be the dual norm function either. Since duality is symmetric, the dual norm of the dual norm is the primal norm and the symmetric normality relations are given in (11). The gradient of the dual norm on the "unit circle" $\|y\|_{(d)} = 1$ must fall on the unit circle (the square $ABCD$) of the primal norm. The diamond, satisfying this condition, is the correct dual unit circle.

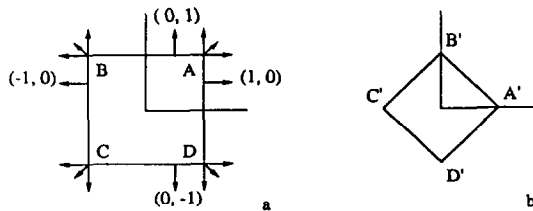


Fig. 3. Duality and normality of 1-norm and ∞ -norm.

The gradient of a C^0 norm function at the vertices is not unique and is interpreted as a point-to-set map. This generalized gradient [5a-c] takes the form of triangular fans as shown in Fig. 3 in which the length of all gradient vectors are proportionally drawn.

In computational approach to the solutions of applied problems, C^0 norms are less popular for the obvious reason of numerical difficulty. A C^0 norm is often replaced by a smooth norm in applications.

APPLICATION TO PLASTICITY

The concept of a norm is not restricted to vectors. There are matrix norms [6], function norms and operator norms [7]. We shall limit our discussion to finite dimensional spaces therefore only vector and matrix norms are discussed. An $m \times n$ matrix can be thought of as an mn -vector. A matrix norm has the same properties as that of the vector norms. It can be either

induced by a vector norm [6] or independently defined. The norms defined on stress matrices in the theory of plasticity are mathematical models of yield behavior based on experimental results and certain physical principles [10a, b].

In plasticity, the natural primal variable is the stress denoted by $\sigma \in R^{3 \times 3}$, a real symmetric 3×3 matrix representing the state of force per unit area at a point in the material. Since there exists no real material that is infinitely strong, the strength of a material can be modeled by a bound

$$\|\sigma\| \leq \sigma_0 \quad (29)$$

on the stress matrix σ where σ_0 is a material constant. The specific form of the norm is called the yield function derived from experimental data. The best known yield function is the von Mises yield function [8]. To shorten the discussion but still keep all the essentials relevant to the present exposition, the von Mises yield function is stated in a subspace of plane stress where an element is a 2×2 symmetric stress matrix

$$\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}, \quad \sigma_{12} = \sigma_{21}, \quad (30)$$

representing a state of stress in a sheet of material. The von Mises yield function can be stated in two equivalent forms

$$\|\sigma\|_v = \sqrt{(\sigma_{11}^2 - \sigma_{11}\sigma_{22} + \sigma_{22}^2 + 3\sigma_{12}^2)} = \sqrt{(\sigma_1^2 - \sigma_1\sigma_2 + \sigma_2^2)}, \quad (31)$$

where σ_1 and σ_2 are the eigenvalues of the plane stress matrix in (30) and $\|\cdot\|_v$ denotes the von Mises norm.

Another yield function named after Tresca [8] is physically more precise but is less used because of its non-smoothness. When applied to the plane stress matrix, it takes the form

$$\|\sigma\|_T = \max\{|\sigma_1|, |\sigma_2|, |\sigma_1 - \sigma_2|\}, \quad (32)$$

where the Tresca norm $\|\sigma\|_T$ can also be expressed in terms of the stress components [8]. One may not find these norms in a mathematics book since they are derived from physical considerations. Nevertheless, they are valid norms or seminorms according to the conditions in (15).

The dual variable in the context of plasticity is the strain rate ϵ , also a 2×2 symmetric matrix for the plane stress case. The inner product (σ, ϵ) of the two matrices appears frequently in the field of mechanics. From the physical principle of nonnegative dissipation, a conclusion [9] was reached that the plastic strain rate matrix is a constant multiple of the gradient of the yield function arranged in the form of a matrix, that is

$$\epsilon = k\nabla\|\sigma\|, \quad (33)$$

where the gradient is taken with respect to each of the stress components in (30). This is known as the normality relation in plasticity. Using the mathematical argument of a sharp upper bound, we reach the same conclusion as that in [9].

A fundamental inequality for the mathematical theory of plasticity is established as

$$|(\sigma, \epsilon)| \leq \|\sigma\|_{(p)}\|\epsilon\|_{(d)} \quad (34)$$

which is the matrix version of the theorem presented in this paper for a general pair of primal and dual norms. Equality holds if the normality relation (33) is satisfied.

For the von Mises and Tresca primal norms, the corresponding dual norms on the plastic strain rate ϵ in terms of its eigenvalues ϵ_1 and ϵ_2 are

$$\begin{aligned} \|\epsilon\|_{\Lambda} &= \frac{2}{\sqrt{3}} \sqrt{(\epsilon_1^2 + \epsilon_1\epsilon_2 + \epsilon_2^2)}, \\ \|\epsilon\|_{\perp} &= \max\{|\epsilon_1|, |\epsilon_2|, |\epsilon_1 + \epsilon_2|\}, \end{aligned} \tag{35}$$

respectively. They are derived from the corresponding primal norms in (31) and (32) by the constructive process described earlier. Since $\|\cdot\|_{\top}$ is a C^0 norm, interpretation of its gradient need be done in a generalized sense for the construction of its dual $\|\cdot\|_{\perp}$.

In [11a], we have introduced a parametric family of (β) -norms for the moment matrix \mathbf{M} in the plate problems

$$\|\mathbf{M}\|_{(\beta)} = \sqrt{(M_1^2 - \beta M_1 M_2 + M_2^2)}, \quad -2 < \beta < 2, \tag{36}$$

in terms of the eigenvalues M_1 and M_2 of \mathbf{M} . It includes the Frobenius norm ($\beta = 0$) [11b] and the von Mises norm ($\beta = 1$) as special cases. By the same construction process, we have shown that the $(-\beta)$ -norm of a curvature rate matrix κ with eigenvalues κ_1 and κ_2 ,

$$\|\kappa\|_{(-\beta)} = \frac{2}{\sqrt{(4 - \beta^2)}} \sqrt{(\kappa_1^2 + \beta\kappa_1\kappa_2 + \kappa_2^2)}, \tag{37}$$

is the appropriate dual norm.

Using these paired non-Euclidean norms in plasticity and the generalized Hölder inequality, we have established several minimax (duality) theorems for certain structural and manufacturing problems. The results can be found in [11a–d]. Since they preceded this paper, the claims to the generalized Hölder inequality were only implicit and its use was less mature. Nevertheless, the patterns emerged from these results inspired the theorem presented in this paper.

A remark

There is a footnote to be added. We discovered another remarkable property of primal–dual relation after this paper was written. Namely, the dual norm of the gradient of the primal norm always equals to unity, that is

$$\|(\nabla\|\mathbf{x}\|_{(p)})\|_{(d)} = 1, \tag{38}$$

for all $\mathbf{x} \in R^n$. The proof can be obtained simply by taking the dual norm of \mathbf{y} in (17). An example

$$\|(\nabla\|\mathbf{x}\|_p)\|_q = 1, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad 1 < p < \infty, \quad \forall \mathbf{x} \in R^n, \quad \mathbf{x} \neq 0$$

can be easily computed.

Acknowledgement—The author is grateful to L. Cesari for his enlightened critic and helpful pointers to some references in the literature of mathematics, leading to substantial improvement of the manuscript. Discussions with B. A. Taylor many years ago seeded the idea.

REFERENCES

1. HARDY G. H., LITTLEWOOD J. E. & POLYA G., *Inequalities*, 2nd edn. Cambridge University Press, Cambridge (1952).
2. HALMOS P. R., *Introduction to Hilbert Spaces*, 2nd edn. Chelsea, NY (1957).

3. BIRKHOFF G. & MACLANE S., *A Survey of Modern Algebra*. Macmillan, NY (1953).
4. GOFFMAN C. & PEDRICK G., *First Course in Functional Analysis*. Prentice-Hall, Englewood Cliffs, NJ (1965).
5. CLARKE F. H., (a) Generalized gradients and applications, *Trans. Am. math. Soc.* **205**, 247–262 (1975); (b) Generalized gradients of Lipschitz functionals, *Adv. Math.* **40**, 52–67 (1981); (c) *Optimization and Nonsmooth Analysis*. John Wiley (1983).
6. HOUSEHOLDER A. S., On norms of vectors and matrices. ORNL report 1759 (1954).
7. ROYDEN H. L., *Real Analysis*, 3rd edn. Macmillan, NY (1988).
8. HILL R., *The Mathematical Theory of Plasticity*. Clarendon Press, Oxford (1950).
9. DRUCKER D. C., A definition of stable inelastic material, *J. appl. Mech.* **81**, (E101) (1959).
10. YANG W. H., (a) A useful theorem for constructing convex yield functions, *J. appl. Mech.* **47**, 301–303 (1980); (b) A generalized von Mises criterion for yield and fracture, *J. appl. Mech.* **47**, 297–300 (1980).
11. YANG W. H., (a) duality theorem for plastic plates, *Acta Mech.* **69**, 177–193 (1987); (b) A variational principle and an algorithm for limit analysis of beams and plates, *Computat. Meth. appl. Mech. Engng.* **33**, 575–582 (1982); (c) Minimization approach to limit solutions of plates, *Computat. Meth. appl. Mech. Engng.* **28**, 265–274 (1981); (d) Pipe flow of plastic materials, *J. appl. Mech.* **47**, 496–498 (1980).