Nonlinear version of Holub's theorem and its application

PENG Jigen and XU Zongben

Research Center for Applied Mathematics, Institute for Information and System Science, Xi'an Jiaotong University, Xi'an 710049, China

Abstract Holub proved that any bounded linear operator T or $-T$ defined on Banach space $L^1(\mu)$ satisfies Daugavet equation

$$1 + \| T \| = \mathrm{Max}\{ \| I + T \|, \| I - T \| \}.$$

Holub's theorem is generalized to the nonlinear case: any nonlinear Lipschitz operator f defined on Banach space l^1 satisfies

$$1 + L(f) = \mathrm{Max}\{L(I + f), L(I - f)\},$$

where $L(f)$ is the Lipschitz constant of f. The generalized Holub theorem has important applications in characterizing the invertibility of nonlinear operator.

Keywords: **nonlinear Lipschitz operator, Holub theorem, Daugavet equation, invertibility of operator.**

LET E be a Banach space. A bounded linear operator $T: E \to E$ is said to satisfy the Daugavet equation if

$$1 + \| T \| = \| I + T \|, \tag{1}$$

where I denotes the identity operator. Daugavet equation is of fundamental importance in many fields, such as the best approximation theory, the geometry of Banach space, and the operator theory. The reseach interest for Daugavet equation has been growing rapidly (for useful references refer to references [1—4]).

Since Daugavet proved that the Daugavet equation is satisfied by any compact operator on $C[0, 1]$, one of the outstanding works on this subject is the following theroem, which belongs essentially to Holub[1, 2].

Holub's theorem. If T is a bounded linear operator on $L^1(\mu)$ (or, in general on an AL- or AM-space), then T satisfies

$$1 + \| T \| = \mathrm{Max}\{ \| I + T \|, \| I - T \| \}; \tag{2}$$

that is, T or $-T$ satisfies Daugavet equation.

Let $f: E \to E$ be a Lipschitz operator. Following ref. [5], the Lipschitz constant and Dalhquist-constant associated with f are defined, respectively, by

$$L(f) = \sup_{x \neq y} \frac{\| f(x) - f(y) \|}{\| x - y \|}, \quad M(f) = \lim_{r \to 0^+} \frac{L(I + rf) - 1}{r}. \tag{3}$$

Obviously, if f is linear, then $L(f) = \| f \|$ (the operator norm of f). Denote by Lip (E) the space of all Lipschitz operators defined on E. It is then easy to show that $L(\cdot)$ is a seminorm of Lip (E) (for details see reference [5]).

The main purpose of the present note is to show that the Holub's theorem can be generalized to every $f \in \mathrm{Lip}(l^1)$ with seminorm $L(\cdot)$ (see Theorem 1 below). We will also apply the generalized theorem to derive some new results on invertibility of nonlinear operator.

Theorem 1. If $f \in \mathrm{Lip}(l^1)$, then f satisfies

$$1 + L(f) = \mathrm{Max}\{L(I + f), L(I - f)\}. \tag{4}$$

In order to prove this theorem, we need to establish the following lemmas.

Lemma 1. If $f \in \mathrm{Lip}(R^n)$, $\| \cdot \|$ is an arbitrary norm on R^n, then for any two points x_1, x_2 in R^n, there exists a matrix A on R^n having the following two properties:

(i) $\| A \| \leqslant L(f)$, $f(x_1) - f(x_2) = A(x_1 - x_2)$;

(ii) $\| A + B \| \leqslant L(f + B)$ for each matrix B on R^n.

Proof. Let L be the line $\{x = tx_1 + (1 - t)x_2 : t \in \mathrm{R}\}$, and Π be the orthogonal subspace $\{x \in R^n : \langle x, y \rangle = 0, y \in L\}$ where $\langle \cdot, \cdot \rangle$ denotes the inner product on R^n. Then, for each $y \in R^n$, there exist uniquely $y_1 \in L$, $y_2 \in \Pi$ such that $y = y_1 + y_2$.

Let $V_m = \{y \in R^n : y = y_1 + y_2, y_1 \in L, y_2 \in \Pi, \| y_2 \| \leqslant m^{-1}$ and $y_1 \in \{tx_1 + (1 - t)x_2 : t \in [0, 1]\}\}$. Since f is Lipschitz-continuous, f is differentiable almost everywhere in V_m by the famous Rademacher theorem. By Fubini theorem, there is $y_m \in V_m$ such that f is differentiable almost everywhere in the segment $E(y_m) = \{x + y_m : x = tx_1 + (1 - t)x_2, t \in [0, 1]\}$; that is, for some subset

$G \subseteqq [0, 1]$ with Lebesgue measure $\mu(G) = 1$, f is differentiable at everywhere in $\{tx_1 + (1-t)x_2 + y_m: t \in G\}$. Set $A_m = \int_G f'(y_m + tx_1 + (1-t)x_2)\mathrm{d}\mu(t)$, we then have

$$\begin{aligned} f(x_1 + y_m) - f(x_2 + y_m) &= \int_G \frac{\mathrm{d}}{\mathrm{d}t} f(y_m + tx_1 + (1-t)x_2)\mathrm{d}\mu(t) \\ &= \int_G f'(y_m + tx_1 + (1-t)x_2)(x_1 - x_2)\mathrm{d}\mu(t) \\ &= A_m(x_1 - x_2). \end{aligned} \tag{5}$$

Since $\| A_m \| \leqslant L(f)$, $\{A_m\}_{m>0}$ has a convergent subsequence $\{A_{m_k}\}_{m_k>0}$, i.e. $A_{m_k} \to A$ $(k \to +\infty)$ for some matrix A. Obviously $\| A \| \leqslant L(f)$, and $f(x_1) - f(x_2) = A(x_1 - x_2)$ follows from (4).

For any matrix B, replacing f by $B + f$ in the above argument shows $\| A + B \| \leqslant L(f + B)$. The proof is therefore completed.

Lemma 2. If $E = (R^n, \| \cdot \|_1)$, then each $f \in \mathrm{Lip}(E)$ satisfies

$$L(f) + 1 = \mathrm{Max}\{ L(I + f), L(I - f)\}. \tag{6}$$

Proof. If $A = (a_{ij})_{n\times n}$ is a real matrix, then by definition $\| A \|_1 = \sup_{\| x \|_1 \leqslant 1} \| Ax \|_1 = \max_{1\leqslant i \leqslant n} \sum_{j=1}^{n} | a_{ij} |$. Without loss of generality, let $\| A \|_1 = \sum_{i=1}^{n} | a_{ik} |$ for some $1 \leqslant k \leqslant n$. If $a_{kk} \geqslant 0$, then

$$\| I + A \|_1 \geqslant | 1 + a_{kk} | + \sum_{i \neq k} | a_{ik} | \geqslant 1 + \| A \|_1.$$

Therefore $\| A + I \|_1 = 1 + \| A \|_1$. If $a_{kk} < 0$, then by a similar way we can conclude $\| I - A \|_1 = 1 + \| A \|_1$.

Let $\varepsilon > 0$ be small sufficiently, and choose $x_\varepsilon, y_\varepsilon \in R^n$ so that $\| f(x_\varepsilon) - f(y_\varepsilon) \|_1 \geqslant (L(f) - \varepsilon) \| x_\varepsilon - y_\varepsilon \|_1$. Then by Lemma 1, there exists a matrix A_ε for x_ε and y_ε such that

$$\begin{aligned} \| f(x_\varepsilon) - f(y_\varepsilon) \|_1 &\leqslant \| A_\varepsilon \|_1 \cdot \| x_\varepsilon - y_\varepsilon \|_1 \\ &\leqslant [\mathrm{Max}\{ \| I + A_\varepsilon \|_1, \| I - A_\varepsilon \|_1\} - 1] \cdot \| x_\varepsilon - y_\varepsilon \|_1 \\ &\leqslant [\mathrm{Max}\{L(f + I), L(I - f)\} - 1] \cdot \| x_\varepsilon - y_\varepsilon \|_1. \end{aligned} \tag{7}$$

Consequently $L(f) + 1 \leqslant \mathrm{Max}\{L(I + f), L(I - f)\} + \varepsilon$, and $L(f) + 1 \leqslant \mathrm{Max}\{L(I + f), L(I - f)\}$ by ε's choice. Since the opposite inequality is obvious, the proof of Lemma 2 is completed.

Proof of Theorem 1. Let $P_m: l^1 \to l^1$, $P_m x = \{x_1, \cdots, x_m, 0, \cdots\}$ (in which $x = \{x_1, \cdots, x_m, x_{m+1}, \cdots\}$). Then $P_m(l^1) \cong (\mathrm{R}^m, \| \cdot \|)$. Denote by h_m the restriction of $P_m f$ on $P_m(l^1)$. So $h_m \in \mathrm{Lip}(R^m)$, and by Lemma 2

$$L(h_m) + 1 = \mathrm{Max}\{L(I_m + h_m), L(I_m - h_m)\} \leqslant \mathrm{Max}\{L(I + f), L(I - f)\}. \tag{8}$$

For any $\varepsilon > 0$, we choose $x, y \in l^1$ such that $\| f(x) - f(y) \| \geqslant (L(f) - \varepsilon) \| x - y \|$, and choose m large enough to satisfy

$$\begin{aligned} &\| P_m(f(x) - f(y)) \| \geqslant (1 - \varepsilon) \| f(x) - f(y) \| \\ &\| x - P_m x \| \leqslant \varepsilon \| x - y \|, \quad \| y - P_m y \| \leqslant \varepsilon \| x - y \| \end{aligned} \tag{9}$$

Then noting that $\| P_m(x - y) \| \leqslant \| x - y \|$, we obtain

$$\begin{aligned} (1 - \varepsilon) \| f(x) - f(y) \| &\leqslant \| P_m f P_m(x) - P_m f P_m(y) \| \\ &\quad + L(P_m f)[\| x - P_m x \| + \| y - P_m y \|] \\ &\leqslant L(h_m) \| P_m x - P_m y \| + 2\varepsilon L(f) \| x - y \|. \end{aligned} \tag{10}$$

That ε is arbitrarily implies $L(f + 1) \leqslant \mathrm{Max}\{L(I + f), L(I - f)\}$ from (8)—(10). The proof is completed.

We now apply Theorem 1 to deriving some results on invertibility of a nonlinear operator.

We first state the following lemma, which is a direct consequence of definition (3) combined with the Banach's contraction mapping theorem.

Lemma 3. Let E be a Banach space and $f \in \mathrm{Lip}(E)$. If $M(f) < 0$, then f is invertible. $f^{-1} \in \mathrm{Lip}(E)$ and $L(f^{-1}) \leqslant -(M(f))^{-1}$.

Suppose $f \in \mathrm{Lip}(E)$. We set $[f] = \{\lambda f: \lambda \in \mathrm{R}\}$ and $\mathrm{dist}(I, [f]) = \inf\{L(I - \lambda f): \lambda \in \mathrm{R}\}$.

If $L(I-\lambda f)=\text{dist}(I,[f])$, then λ is said to be the best approximation of I to $[f]$. Obviously $\text{dist}(I,[f])\leqslant 1$. If $\text{dist}(I,[f])<1$, then by Lemma 3, f is invertible and $f^{-1}\in \text{Lip}(E)$, and hence the invertibility of f is tightly related to the best approximation of I to $[f]$.

Lemma 4. Suppose that E is a Banach space and $f\in \text{Lip}(E)$. Then at least one of the following propositions holds:

(i) f is invertible and $f^{-1}\in \text{Lip}(E)$;

(ii) $\lambda=0$ is the unique best approximation of I to $[f]$; that is, $L(I-\lambda f)>1$ for all $\lambda\neq 0$;

(iii) for all $\lambda>0$, $\lambda I+f$ is invertible and $L[(\lambda I+f)^{-1}]=\lambda^{-1}$;

(iv) for all $\lambda>0$, $\lambda I-f$ is invertible and $L[(\lambda I-f)^{-1}]=\lambda^{-1}$.

Proof. Suppose that both (i) and (ii) are wrong. Then f is not invertible. And there exists $\lambda_0\neq 0$ such that $L(I-\lambda_0 f)=1$. In this case, for each $\alpha\geqslant 1$, from the relation $1=L(I-\lambda_0 f)=L(\alpha I-\lambda_0 f+(1-\alpha)I)\geqslant L(\alpha I-\lambda_0 f)-(\alpha-1)$, we have $L(I-\alpha^{-1}\lambda_0 f)=1$ (note that f is not invertible). If $\lambda_0>0$, then by the choice of α, $L(I-tf)=1$ for all $t\in[0,\lambda_0]$, and $M(-f)=0$ by formula (3). Because $M(-\lambda I-f)=-\lambda+M(-f)=-\lambda<0$ for each $\lambda>0$[5], so by Lemma 3 $\lambda I+f$ is invertible and $L[(\lambda I+f)^{-1}]\leqslant\lambda^{-1}$. If $L[(\lambda I+f)^{-1}]<\lambda^{-1}$, then $f\circ(\lambda I+f)^{-1}(=1-\lambda(\lambda I+f)^{-1})$ is invertible. This contradicts the fact that f is not invertible, and therefore, proposition (iii) holds.

If $\lambda_0<0$, we can show similarly that proposition (iv) holds. This completes the proof.

Theorem 2. Suppose that $f\in \text{Lip}(l^1)$ and $L(f)=1$. Then at least one of the following propositions holds:

(i) f is invertible and $f\in \text{Lip}(l^1)$;

(ii) for each $0\neq\lambda\in(-2,2)$, $l(I+\lambda f)\triangleq\inf\limits_{x\neq y}\dfrac{\|(I+\lambda f)(x)-(I+\lambda f)(y)\|}{\|x-y\|}\neq 1$ or the range $\overline{R(I+\lambda f)}\neq l^1$;

(iii) there exists an $\varepsilon\neq 0$ such that $I+\lambda f$ is invertible and $(I+\lambda f)^{-1}\in \text{Lip}(l^1)$ either for all $\lambda\geqslant\varepsilon$ or for all $\lambda\leqslant\varepsilon$.

Proof. Suppose that f is not invertible and there exists $0\neq\varepsilon\in(-2,2)$ such that $l(I+\varepsilon f)=1$ and $\overline{R(I+\varepsilon f)}=l^1$. It is easy to show that $I+\varepsilon f$ is invertible and $L[(I+\varepsilon f)^{-1}]=1$. So by Theorem 1 there is nothing more than the following two cases.

Case 1. $L[I+(I+\varepsilon f)^{-1}]=1+L[(I+\varepsilon f)^{-1}]=2$. In this case, $L\left[I-\frac{\varepsilon}{2}f(I+\varepsilon f)^{-1}\right]=1$. Since $f(I+\varepsilon f)^{-1}$ is not invertible, by Lemma 4 we have that either $[\lambda I+(1+\lambda\varepsilon)f)](I+\varepsilon f)^{-1}(=\lambda I+f(I+\varepsilon f)^{-1})$ is invertible in l^1 for all $\lambda>0$, or $[\lambda I+(\lambda\varepsilon-1)f)](I+\varepsilon f)^{-1}(=\lambda I-f(I+\varepsilon f)^{-1})$ is invertible in l^1 for all $\lambda>0$. Therefore proposition (3) holds by λ's choice.

Case 2. $L[I-(I+\varepsilon f)^{-1}]=1+L[(I+\varepsilon f)^{-1}]=2$. In this cases, $2=L[\varepsilon f(I+\varepsilon f)^{-1}]\leqslant L(\varepsilon f)\cdot L[(I+\varepsilon f)^{-1})]=L(\varepsilon f)$, and hence $|\varepsilon|\geqslant 2$, which contradicts $\varepsilon\in(-2,2)$. The proof is complete.

Acknowledgement Thiw work was supported by Doctorate Foundation of Xi'an Jiaotong University.

References

1 Holub, J. R., Daugavet's equation and operators on $L^1(\mu)$, *Proc. Amer. Math. Soc.*, 1987, 100: 295.

2 Schmidt, K. D., Daugavet's equation and orthomorphisms, *Proc. Amer. Math. Soc.*, 1990, 108(4): 905.

3 Weis, L., An elementary approach to Daugavet equation, in *Lect. Notes in Pure Appl. Math.* 175, New York and Basel: Marcel Dekker Inc., 1996, 494.

4 Holub, J. R., A note on best approximation and invertibility of operators on uniformly convex banach space, *Internat. J. Math. and Math. Sci.*, 1991, 14(3): 611.

5 Soderlind, G., Bounds on nonlinear operators in finite-dimensional Banach space, *Numer. Math.*, 1986, 50: 27.

(*Received May* 7, 1997)