03nov99

III. Normed linear spaces

** definition **

$(X, \|\cdot\|)$ is a **normed linear space** (=:**nls**) := X is a ls and $\|\cdot\|$ is a **norm** on X, i.e., $\|\cdot\| : X \to \mathbb{R}$ and satisfies

$$\begin{aligned} \|x\| &\ge 0, \qquad \text{with equality iff } x = 0 && \textbf{(positive definite)} \\ \|\alpha x\| &= |\alpha|\|x\| && \textbf{(absolute homogeneous)} \\ \|x+y\| &\le \|x\| + \|y\| && \textbf{(subadditive)} \end{aligned}$$

The last inequality is the **triangle inequality**. *Examples* are provided by the earlier samples $\ell_p(m) := (\mathbb{R}^m, \|\cdot\|_p)$, $(C([a \mathinner{..} b]), \|\cdot\|_p)$ and $C[a \mathinner{..} b] := \big(C([a \mathinner{..} b]), \|\cdot\|_\infty\big)$, but many more will follow.

Any nls is a ts, in particular a ms, with the **norm metric** given by

$$d(x,y) := d_{\|\|}(x,y) := \|x-y\|.$$

The norm is continuous (in fact Lipschitz continuous with constant 1) with respect to this metric, as can be seen by substituting $x-y$ for x in the triangle inequality, getting $\|x\| - \|y\| \le \|x-y\|$, hence, by symmetry,

$$|\ \|x\| - \|y\|\ | \le \|x-y\|.$$

H.P.(1) A **seminorm** $\|\cdot\|$ only lacks definiteness to be a norm, i.e., it has all the attributes of a norm except that $\|x\|$ may be 0 without x being 0. Assuming $(X, \|\cdot\|)$ to be a *semi*normed ls, prove:
(i) $Y := \ker \|\cdot\| = \{x \in X : \|x\| = 0\}$ is a lss.
(ii) $X/Y = \{\langle x\rangle := x + Y : x \in X\}$ is a nls with the induced norm $\|\langle x\rangle\| := \|x\|$.
In other words, a seminormed ls can always be considered a nls if one is prepared to consider x and y to be the same in case their difference has (semi)norm 0.

H.P.(2) Verify that $\|f\| := \limsup_{n\to\infty} f(n) - \liminf_{n\to\infty} f(n)$ is a seminorm on the ls $m := b(\mathbb{N})$ of all bounded sequences. Conclude that $c :=$ collection of all convergent sequences is a lss of m. Can you think of a 'natural' lm on m whose kernel is c? (I can't.)

** norm metric **

The most important feature of a norm metric (when compared to other metrics) is its **translation invariance**:

$$d(x,y) = d(x+z, y+z).$$

This implies that the nbhdsystems for any two points are alike,

$$\mathbf{B}(x) = \mathbf{B}(0) + x,$$

and so allows to settle many questions by considering just neighborhoods of 0,

$$B_r := B_r(0).$$

For example, if X, Y are nls's and $f \in L(X,Y)$, then $f(B_r(x)) = f(B_r) + f(x)$, therefore

$$f(B_r(x)) \subseteq B_s(f(x)) \iff f(B_r) + f(x) \subseteq B_s + f(x)$$
$$\iff f(B_r(0)) \subseteq B_s(f(0))$$

(using the fact that, by additivity, $0 = f(0)$). Thus, $f \in L(X,Y)$ is continuous at x iff f is continuous at 0. Since x is arbitrary here, this says that $f \in L(X,Y)$ is continuous at a point iff f is continuous (everywhere).

The second most important feature of a norm metric is its **scale invariance**:

$$d(\alpha x, \alpha y) = |\alpha| d(x,y).$$

This implies that the entire nbhdsystem $\mathbf{B}(0) = \{B_r : r > 0\}$ is obtainable from just one nbhd by scaling,

$$B_r = rB_1, \qquad \text{all } r > 0.$$

This means that we can understand $\mathbf{B}(0)$ by understanding the **unit ball**

$$B_X := B := B_1 = B_1(0).$$

H.P.(3) Y lss of nls X. Prove: *For any $x \in X$, $y \in Y$, $d(x+y, Y) = d(x, Y)$.*

H.P.(4) Y lss of nls X. Prove: $X \to \mathbb{R} : x \mapsto d(x,Y)$ *is a seminorm*, hence conclude that X/Y *is nls wrto* $\langle x \rangle \mapsto d(x,Y)$ *iff* Y *is closed.* (Recall from (II.10) that $Y^- = \{x \in X : d(x,Y) = 0\}$.)

H.P.(5) Prove: *Any* **linear** *(:= translation- and scale-invariant) metric is a norm metric.*

A ls X with a metric that is only translation-invariant but not scale-invariant is called a **Fréchet space** if it is also complete and if the map $X \times \mathbb{F} : (x, \alpha) \mapsto x\alpha$ is continuous in each of its two arguments separately. (It then can be shown that it is necessarily continuous as a map from $X \times \mathbb{F}$ to X.)

H.P.(6) Prove that the space $X = \mathbb{C}^{\mathbb{Z}^n}$ of all complex-valued functions on the set $\mathbb{Z}^n$ is a Fréchet space with respect to the topology of pointwise convergence.

It can be shown that the topology on any Fréchet space is, equivalently, that of convergence with respect to an at most countable set of seminorms.

** boundedness and continuity **

Let X, Y be nls's. We say that $f \in L(X,Y)$ *is* **bounded** if it carries bounded sets to bounded sets. (Warning: This is at variance with the standard definition according to which a map, linear or not, is bounded if its range is bounded, but what can I do?) The scale invariance also implies that *a continuous linear map is bounded*: Indeed, if $U \subseteq X$ is bounded, then $U \subseteq B_r(x)$ for some x and r, hence $U \subseteq B_s$ for some s (e.g., $s = r + \|x\|$). If f is also continuous, then there exists $t > 0$ so that $f(B_t) \subseteq B$. But then, $f(U) \subseteq f(B_s) = f((s/t)B_t) = (s/t)f(B_t) \subseteq B_{s/t}$, i.e., $f(U)$ is bounded.

Conversely, *a bounded $f \in L(X,Y)$ is continuous*: If $f(B)$ is bounded, then $f(B) \subseteq B_s$ for some s. But then, for any $r > 0$, $f(B_{r/s}) \subseteq B_r$, i.e., f is continuous at 0, hence continuous.

Thus, for a lm, continuity and boundedness are one and the same:

(1) Proposition. *$f \in L(X, Y)$ is continuous iff $f(B)$ is bounded.*

(Aside: Since boundedness says that $\forall\{r > 0\}\exists\{s > 0\}\ f(B_r) \subseteq B_s$, while continuity says that $\forall\{s > 0\}\exists\{r > 0\}\ f(B_r) \subseteq B_s$, the equivalence between boundedness and continuity is a kind of saddle point.)

H.P.(7) Prove: *If $f \in L(X, Y)$ maps some nonempty open set to a bounded set, then f is continuous.*

We denote by

$$bL(X, Y)$$

the collection of all bounded linear maps from the nls X to the nls Y.

H.P.(8) Prove: *$bL(X, Y)$ is a linear space* (as a linear subspace of $L(X, Y)$).

(2) Proposition. *$bL(X, Y)$ is a normed linear space, with respect to the* **map norm**

$$(3)\qquad \begin{aligned} \|f\| := \|f : X \to Y\| &:= \sup \|f(B)\| := \sup_{x \in X}\{\|f(x)\| : \|x\| < 1\} = \sup \|f(B^-)\| = \\ &= \sup\{\|f(x)\| : \|x\| = 1\} = \sup_{x \neq 0} \|f(x)\|/\|x\| = \min\{r : f(B) \subseteq B_r^-\} \end{aligned}$$

defined for every $f \in bL(X, Y)$ (since $f(B)$ is bounded for such f).

Proof: If $f \in bL(X, Y)$, then f is continuous, hence so is $X \to \mathbb{R} : x \mapsto \|f(x)\|$, and therefore $\sup \|f(B)\| = \sup \|f(B^-)\|$ (by (II.8)Proposition). Further, for $x \neq 0$, $\|f(x)\| = \|x\|\|f(x/\|x\|)\|$, hence $\sup \|f(B^-)\| = \sup \|f(S)\|$, with

$$S := S_X := \{x \in X : \|x\| = 1\}$$

the **unit sphere** for X. But, this also says that $\sup \|f(S)\| = \sup_{x \in X\backslash 0} \|f(x)\|/\|x\|$. Finally, for any $Z \subseteq \mathbb{R}$, $\sup Z$ is the least upper bound for Z, i.e., equals $\min\{r \in \mathbb{R} : \forall\{z \in Z\}\ z \leq r\}$, hence the equality $\sup \|f(B)\| = \min\{r : f(B) \subseteq B_r^-\}$ follows.

Now we verify that the map $bL(X, Y) \to \mathbb{R}_+ : f \mapsto \|f\|$ is a norm. First, $\|f\| = 0$ implies that $\|f(x)\| = 0$ for all x, therefore $f = 0$. Also, $\|(\alpha f)(x)\| = |\alpha|\|f(x)\|$, hence $\|\alpha f\| = |\alpha|\|f\|$. Finally, $\|(f + g)(x)\| = \|f(x) + g(x)\| \leq \|f(x)\| + \|g(x)\|$; therefore, $\|f + g\| = \sup_x \|(f + g)(x)\|/\|x\| \leq \sup_x (\|f(x)\|/\|x\| + \|g(x)\|/\|x\|) \leq \|f\| + \|g\|$. □

H.P.(9) Prove: *The composition of bounded lm's is bounded. More precisely,*

$$(4)\qquad \forall\{f \in bL(X, Y), g \in bL(Y, Z)\}\quad \|gf\| \leq \|g\|\|f\|.$$

** bounded below **

The map $A \in L(X, Y)$ is *bounded* in case, for some finite M and all $x \in X$, $\|Ax\| \leq M\|x\|$, in which case M is **a bound for** A. We call $A \in L(X, Y)$ **bounded below** in case, for some *positive* m and all $x \in X$, $\|Ax\| \geq m\|x\|$, in which case m is **a lower bound for** A. Just as A is bounded precisely when $\|A\| = \sup_{x \in X} \|Ax\|/\|x\|$ is finite, so A is bounded below precisely when $\inf_{x \in X} \|Ax\|/\|x\|$ is positive.

If A is invertible, then we recognize in this infimum the reciprocal of the supremum that gives the norm of A^{-1}. Indeed, assuming $A \in L(X,Y)$ to be invertible,

$$\sup_{y\in Y} \frac{\|A^{-1}y\|}{\|y\|} = \sup_{x\in X} \frac{\|A^{-1}Ax\|}{\|Ax\|} = 1/\inf_{x\in X} \bigl(\|Ax\|/\|x\|\bigr).$$

Therefore, an invertible lm is boundedly invertible (i.e., has a *bounded* inverse) iff it is bounded below and, in that case, we get

$$\|Ax\| \ge \|x\|/\|A^{-1}\|,$$

and this bound is **sharp** since, for any $r > 1/\|A^{-1}\|$ there is some $x \in X$ with $\|Ax\| < r\|x\|$.

A lower bound provides a quantification of being 1-1, but it does not imply invertibility since a lm may be bounded below without being onto.

H.P.(10) Prove: *Any onto lm that is bounded below is* **open**, *i.e., carries open sets to open sets.*

H.P.(11) Prove: *If X is complete and $A \in bL(X,Y)$ is bounded below, then* $\operatorname{ran} A$ *is closed.*

**** any lm on a finite-dimensional domain is continuous ****

An important *example* is provided by $L(X,Y)$ in case $\dim X < \infty$.

(5) Proposition. *X, Y nls's, $\dim X < \infty$. Then $L(X,Y) = bL(X,Y)$.*

Proof: The proof consists in factoring $f \in L(X,Y)$ as $(fV)V^{-1}$, with V a basis for X, and proving that the two lm's $g := fV$ and V^{-1} are continuous.

$$\begin{array}{ccc} X & \xrightarrow{\;f\;} & Y \\ V\nwarrow & & \nearrow g \\ & \mathbb{F}^m & \end{array}$$

(i) Any $g \in L(\mathbb{F}^m, Y)$ is of the form $g = [g_1, g_2, \ldots, g_m]$, hence $\|ga\| = \|\sum_j g_j a(j)\| \le \sum_j \|g_j\|\|a\|_\infty$, and therefore $\|g : \ell_\infty(m) \to Y\| \le \sum_{j=1}^m \|g_j\|$, with each $\|g_j\| < \infty$ (since it is the norm of some element of the nls Y), and the sum involving only finitely many such terms. Conclusion: $L(\mathbb{F}^m, Y) = bL(\ell_\infty(m), Y)$.

(ii) Since $m := \dim X < \infty$, there exists $V \in L(\mathbb{F}^m, X)$ 1-1 and onto. By (i), V is continuous (as a map from $\ell_\infty(m)$), hence so is $a \mapsto \|Va\|$, and this function therefore takes on its infimum on any closed and bounded set in $\mathbb{F}^m$. In particular, there exists b so that

$$\|b\| = 1 \quad \text{and} \quad \|Vb\| = \inf \|V(S)\|$$

with

$$S := \{a \in \mathbb{F}^m : \|a\|_\infty = 1\}$$

the **unit sphere** (which is closed as the preimage of the closed set $\{1\}$ under the continuous map $a \mapsto \|a\|_\infty$). Since V is 1-1 and $b \ne 0$, we have $Vb \ne 0$. This shows that

$$\inf_a \|Va\|/\|a\|_\infty = \inf_a \|V(a/\|a\|_\infty)\| = \inf \|V(S)\| = \|Vb\| > 0,$$

i.e., V is bounded below. But this says that V^{-1} is bounded, hence continuous.

(iii) If $f \in L(X,Y)$, then $f = (fV)V^{-1}$ and $fV \in L(\mathbb{F}^m, Y) = bL(\mathbb{F}^m, Y)$, by (i) while $V^{-1} \in bL(X, \mathbb{F}^m)$ by (ii), so $f \in bL(X,Y)$. □

It takes an infinite-dimensional nls to support an unbounded lm. Even on an infinite-dimensional nls, it is hard to construct an unbounded lm unless the space is *incomplete*. A standard *example* is differentiation as a map on $C^{(1)}$ equipped with the max-norm. There are smooth functions of max-norm ≤ 1 but with arbitrarily large first derivative, e.g., the functions $x \mapsto \sin(nx)$ for 'large' n. Note that $(C^{(1)}, \|\cdot\|_\infty)$ is not complete since the uniform limit of differentiable functions need not be differentiable.

**** closed & bounded = compact iff finite-dimensional ****

(6) Corollary. *Any two norm metrics on a finite-dimensional ls are equivalent.*

Proof: If X is nls and $\dim X < \infty$ and $Y = (X, \|\cdot\|')$, then, by (5)Prop., the identity map is in $bL(X,Y)$ as well as in $bL(Y,X)$. In terms of bounds, this says that, for any two norms $\|\cdot\|, \|\cdot\|'$ on a finite-dimensional ls X, there exist positive constants m, M so that

$$\forall\{x \in X\} \quad m\|x\| \le \|x\|' \le M\|x\|.$$

□

For *example*, on $\mathbb{R}^m$,

$$m^{-1/2}\|x\|_1 \le \|x\|_2 \le \|x\|_1$$

and

$$\|x\|_\infty \le \|x\|_2 \le m^{1/2}\|x\|_\infty.$$

This is related to the fact that, on any finite-dimensional nls, bounded and closed sets are compact, since, with $V : \mathbb{R}^m \to X$ any 1-1 onto linear map and Y any closed and bounded set in X, $V^{-1}(Y)$ is closed and bounded in $\ell_\infty(m)$, hence compact, hence so is $Y = V(V^{-1}(Y))$.

H.P.(12) Prove: *The closed unit ball of a finite-dimensional lss of a nls X is compact, hence closed (in X), and therefore any finite-dimensional lss is closed (in X).*

H.P.(13) Y, Z lss's of nls X. Prove: Y *closed* & $\dim Z < \infty \implies Y + Z$ *closed.* (Hint: $Y + Z$ is the inverse image of $\langle\rangle(Z)$ under $\langle\rangle : X \to X/Y$.)

H.P.(14) Prove: *Any finite-dimensional lss of a nls X provides ba's to any $x \in X$.*

In contrast, the unit ball in any *infinite*-dimensional nls fails to be totally bounded, as follows from (9)Corollary of

(7) Riesz' Lemma. *For any nondense lss Y of a nls X, $\sup_{x\in X} \frac{d(x,Y)}{\|x\|} = 1$.*

Proof:

$$\sup_{x\in X\backslash 0} \frac{d(x,Y)}{\|x\|} = \sup_{z\in X\backslash Y^-} \sup_{y\in Y} \frac{d(z-y,Y)}{\|z-y\|} = \sup_{z\in X\backslash Y^-} \frac{d(z,Y)}{\inf_{y\in Y}\|z-y\|} = \sup_{z\in X\backslash Y^-} \frac{d(z,Y)}{d(z,Y)} = 1.$$

□

H.P.(15) Where in this proof is the fact used that Y is not dense?

Note that Riesz' Lemma is much more than the trivial fact that $d(x,Y) \le \|x\|$. It is equivalent to the following

(8) Corollary. *For any proper closed lss Y of the nls X, the quotient map $\langle\rangle : X \to X/Y$ has norm 1.*

(9) Corollary. *If U is a finite ε-net for B_X for some $\varepsilon < 1$, then $\dim X \le \#U$.*

Proof: For all $x \in X\backslash 0$, $x\sqrt{\varepsilon}/\|x\| \in B_X$, hence there exists $u \in U$ so that $\|x\sqrt{\varepsilon}/\|x\| - u\| < \varepsilon$. This implies that $d(x, \mathrm{ran}[U])/\|x\| < \sqrt{\varepsilon} < 1$, hence, by (7)Lemma, $\mathrm{ran}[U]$ is dense in X. Since $\mathrm{ran}[U]$ is closed by H.P.(12), $X = \mathrm{ran}[U]$. □

H.P.(16) Prove: *The unit ball in any infinite-dimensional nls fails to be totally bounded.*

H.P.(17) Use the same kind of argument as in the proof of Riesz' Lemma to prove:

$$\forall\{A \in bL(X,Y)\backslash 0\}\ \|A\| = \sup \|Ax\|/d(x, \ker A).$$

Conclude that the factor map $A_| : X/\ker A \to Y : \langle x\rangle \mapsto Ax$ has the same norm as A.

Here is another way of stating Riesz' Lemma. Let Y be a lss of the nls X. Then $d(\alpha x, Y) = |\alpha| d(x,Y)$, hence $\sup_{x\in B_r} d(x,Y) = \sup_{x\in B_1} d(rx,Y) = r\sup_{x\in B_1} d(x,Y)$, while, by (II.8)Proposition and the continuity of $x \mapsto d(x,Y)$,

$$\sup_{\|x\|<1} d(x,Y) = \sup_{\|x\|\le 1} d(x,Y) = \sup_{\|x\|=1} d(x,Y) = \sup_x d(x/\|x\|, Y) = \sup_x d(x,Y)/\|x\|.$$

Thus Riesz' Lemma states that $\sup_{x\in B_r} d(x,Y) = r$ in case Y is a nondense lss. In other words: *If $B_r \subseteq B_s(Y)$ for some proper closed lss Y, then $r \le s$.* More generally, we have the following.

(10) Corollary. *$B_r(x) \subseteq B_s(Y)$ for some proper closed lss $Y \implies r \le s$.*

Proof: $B_r \subseteq \big(B_r(x) + B_r(-x)\big)/2 \subseteq \big(B_s(Y) - B_s(Y)\big)/2 \subseteq B_s(Y)$, hence $r \le s$. (Here we have used the facts that, since Y is a lss, $B_s(Y) = -B_s(Y)$ and $B_{2s}(Y)/2 = B_s(Y)$, while, for any set M, $B_t(-M) = -B_t(M)$. Also, since $d(\cdot, Y)$ is a seminorm, we have $B_t(Y) + B_u(Y) \subset B_{t+u}(Y)$.) □

This is as convenient a place as any to prove the following proposition of use later.

(11) Proposition. *If X is a complete nls and Y a closed lss, then X/Y is complete (with respect to its 'natural' norm $\|\langle x\rangle\| := d(x,Y)$).*

Proof: By H.P.(4), X/Y is a nls with respect to the **factor norm** $\langle x\rangle \mapsto d(x,Y)$, and the factor map $x \mapsto \langle x\rangle$ is continuous (in fact, it has norm 1 if Y is a proper lss, by (8)). To show that X/Y is complete, it is, by H.P.(II.25), sufficient to prove that any Cauchy sequence in X/Y has limit points. For this, let $(\langle x_n\rangle)$ be a Cauchy sequence. Then $\lim_{n\to\infty} \mathrm{diam}\{\langle x_j\rangle : j \ge n\} = 0$, hence there exists a strictly increasing $\mu : \mathbb{N} \to \mathbb{N}$ so that $\mathrm{diam}\{\langle x_j\rangle : j \ge \mu(n)\} < 2^{-n}$, all n. This implies the existence of (y_n) in Y so that $\|x_{\mu(n+1)} - x_{\mu(n)} - y_n\| < 2^{-n}$. Therefore, with $z_n := x_{\mu(n)} - \sum_{j<n} y_j$, all n, we have $\|z_{n+1} - z_n\| < 2^{-n}$, hence $\mathrm{diam}(z_{\ge n}) < \sum_{j\ge n} 2^{-j} = 2^{1-n}$, showing that (z_n) is Cauchy, hence has a limit, z say, since X is complete. But then $\langle z\rangle = \lim_n \langle z_n\rangle = \lim_n \langle x_{\mu(n)}\rangle$. □

**** computing the norm of a lm ****

The norm $\|A\|$ of $A \in bL(X,Y)$ is the smallest M so that $\|Ax\| \le M\|x\|$, all $x \in X$. Determination of $\|A\|$ involves two steps:

(i) Show that $\forall\{x \in X\}\ \|Ax\| \le M\|x\|$. Then $\|A\| \le M$, i.e., M is an **upper bound**.

(ii) Show that no $M' < M$ will do. This is usually done by examining closely the string of inequalities used in the proof that M is an upper bound, with the aim of showing that every one of these inequalities is **sharp**, i.e., can be made arbitrarily close to an equality by an appropriate choice of x. This step is easiest if one can find $x \in X\backslash 0$ s.t. $\|Ax\| = M\|x\|$, since this implies that $\|A\| \ge M$, and so, in conjunction with (i), proves that A **takes on its norm at** x, i.e.,

$$\|Ax\| = \|A\|\|x\|.$$

If $\dim X < \infty$, then every $A \in bL(X,Y)$ takes on its norm. Indeed, $x \mapsto \|Ax\|$ is continuous, hences takes on its supremum on the closed unit ball B^-, since B^- is compact when $\dim X < \infty$.

(12) Example. For $A \in L(\mathbb{F}^n, \mathbb{F}^m) = \mathbb{F}^{m\times n}$,

$$\begin{aligned}
\|A\|_1 &:= \sup \|Ax\|_1/\|x\|_1 &= \max_j \textstyle\sum_i |A(i,j)| \\
\|A\|_\infty &:= \sup \|Ax\|_\infty/\|x\|_\infty &= \max_i \textstyle\sum_j |A(i,j)| \\
\|A\|_{1,\infty} &:= \sup \|Ax\|_\infty/\|x\|_1 &= \max_{i,j} |A(i,j)|
\end{aligned}$$

(as you should verify!). There is no nice formula for $\|A\|_p := \sup \|Ax\|_p/\|x\|_p$ for $1 < p < \infty$ since in this case the unit ball in $\mathbb{F}^n$ has infinitely many extreme points (cf. Chapter VI: Convexity). The next best thing is $\|A\|_2$ which equals the squareroot of the largest eigenvalue of $A'A$.

If $\dim X \not< \infty$, then $A \in bL(X,Y)$ need not take on its norm. Here is a simple example.

(13) Example. For

$$X = \ell_1 := \{x \in \mathbb{R}^{\mathbb{N}} : \|x\|_1 := \sum |x(i)| < \infty\},$$

let

$$\lambda : X \to \mathbb{R} : x \mapsto \sum_i (1 - 1/i)x(i).$$

Then $|\lambda x| \le \sum_i |1 - 1/i||x(i)| \le \sum_i |x(i)| = \|x\|_1$, so $\|\lambda\| \le 1$, while $\|\lambda\| = \|\lambda\|\|e_j\|_1 \ge |\lambda e_j| = |1 - \frac{1}{j}| \xrightarrow[j\to\infty]{} 1$, so $\|\lambda\| \ge 1$. So, $\|\lambda\| = 1$. But, for any i, $|1 - 1/i||x(i)| = |x(i)|$ iff $x(i) = 0$, thus $|\lambda x| = \|\lambda\|\|x\|_1$ implies $\|x\|_1 = 0$.

An application of map norm: Approximate inverse

$A \in bL(X,Y), C \in bL(Y,X)$. Call C an **approximate (left) inverse** for A in case $\|1 - CA\| < 1$.

H.P.(18) Give an example to show that $\|1 - CA\| < 1$ does not imply that $\|1 - AC\| < 1$ even if $X = Y$.

With
$$E := 1 - CA,$$
an approximate inverse supplies a **lower bound** for A: Since $CAx = x - Ex$, one gets
$$\|C\|\|Ax\| \geq \|CAx\| \geq \|x\| - \|Ex\| \geq (1 - \|E\|)\|x\|,$$
or
$$\|Ax\| \geq \frac{1 - \|E\|}{\|C\|}\|x\|.$$
With x the solution to the linear equation $A? = y$, this provides the only realistic *a posteriori* way to bound the error $x - \tilde{x}$ in an approximate solution $\tilde{x}$ in terms of the computable **residual** $r := y - A\tilde{x} = A(x - \tilde{x})$, viz. the bound
$$\|x - \tilde{x}\| \leq \frac{\|C\|}{1 - \|E\|}\|r\|.$$

More than that, it suggests that one *correct* the current guess $\tilde{x}$ for the solution x by adding to it the computable error 'estimate' $Cr = C(y - A\tilde{x}) = CA(x - \tilde{x}) \sim x - \tilde{x}$. This is at the basis of successful *fixed point iteration*: The equation $Ax = y$ implies the equation
$$x = x + C(y - Ax) = Ex + Cy,$$
and, assuming X to be complete, the fact that $\|E\| < 1$ ensures convergence of the iteration

(14) $$x_{n+1} := Ex_n + z, \qquad n = 0, 1, 2, \ldots,$$

(regardles of the choice of $z \in X$) since it ensures that the iteration function $g : x \mapsto Ex + z$ is Lipschitz continuous with Lipschitz constant $\kappa = \|E\| < 1$. Since x is the (unique) fixed point of (14) iff $CAx = z$, it follows that the linear equation $CA? = z$ has exactly one solution for every $z \in X$. This means that CA is invertible. Further, $(CA)^{-1}$ is bounded since CA is bounded below: $\|CAx\| \geq (1 - \|E\|)\|x\|$, hence
$$\|(CA)^{-1}\| \leq 1/(1 - \|E\|).$$
Finally, since
$$x_n = Ex_{n-1} + z = E^2 x_{n-2} + Ez + z = \cdots = E^n x_0 + \sum_0^{n-1} E^j z,$$
we have, with $x_0 = 0$, that $x := (CA)^{-1}z = \lim \sum_{j<n} E^j z$. In fact, from (II.21),
$$\|(CA)^{-1}z - \sum_{j<n} E^j z\| = d(x, x_n) \leq \underbrace{d(x_1, x_0)}_{= \|z\|} \frac{\|E\|^n}{1 - \|E\|},$$
hence
$$\|(CA)^{-1} - \sum_{j<n} E^j\| \leq \|E\|^n/(1 - \|E\|).$$

This proves

(15) Proposition. *If $A \in bL(X,Y)$ has the approximate inverse $C \in bL(Y,X)$, and X is complete, then CA is boundedly invertible, and the inverse is representable as a* **Neumann series**,

$$(CA)^{-1} = \sum_{n=0}^{\infty} E^n \qquad \textit{(with the limit taken in the norm metric)},$$

with $E := 1 - CA$. Further, $(CA)^{-1}C$ is a (bounded) left inverse for A, and $A(CA)^{-1}$ is a (bounded) right inverse for C. Hence, if C is 1-1 or A is onto, then both A and C are (boundedly) invertible and $A^{-1} = (CA)^{-1}C, \quad C^{-1} = A(CA)^{-1}$.

This proposition is at the basis of most arguments for the bounded invertibility of a linear map.

To substantiate the last claim of (15), observe that the invertibility of CA implies that C is onto and A is 1-1. Hence, if, e.g., C is also 1-1, then it is invertible, therefore $A = C^{-1}(CA)$ or $A^{-1} = (CA)^{-1}C$, showing that A is boundedly invertible, and $C^{-1} = A(CA)^{-1}$ also bounded.

Remark. If $\dim Y \le \dim X < \infty$, then the invertibility of CA implies the invertibility of C and A, but otherwise neither C nor A need be invertible. In particular, the original equation $A? = y$ need not be solvable. (E.g., if A is the right shift on $X = \ell_\infty := b(\mathbb{N})$, i.e.,

$$Ax : j \mapsto \begin{cases} x(j-1), & j > 1; \\ 0, & j = 1, \end{cases},$$

then $CA = 1$ for C given by $Cx : j \mapsto x(j+1)$, all j, but the equation $A? = e_1$ is not solvable.) We only conclude that the equation $CA? = Cy$, in fact the equation $CA? = z$ for arbitrary z, has a unique solution.

H.P.(19) Prove that, for a complete nls X and boundedly invertible $A \in bL(X,Y)$,

(16) $$d(A, \{Q \in bL(X,Y) : Q \text{ not boundedly invertible}\}) \ge 1/\|A^{-1}\|.$$

(Hint: Show that A^{-1} is an approximate inverse for any Q with $\|A - Q\| < 1/\|A^{-1}\|$.) Conclude that $\{A \in bL(X,Y) : A \text{ is boundedly invertible}\}$ *is open*.

Actually, equality must hold in (16), as can be seen by constructing an appropriate rank-one modification of A, as in the proof of the following formal statement.

(17) Proposition. *If X is complete and $A \in bL(X,Y)$ is boundedly invertible, then*

$$d(A, \{Q \in bL(X,Y) : Q \text{ not boundedly invertible}\}) = 1/\|A^{-1}\|.$$

Proof: The (bounded) rank-one linear maps from X to Y are of the form

$$[y]\lambda : X \to Y : x \mapsto y\,(\lambda x), \qquad \forall \{y \in Y, \lambda \in bL(X, \mathbb{F})\},$$

and the norm of such a map is just the product of the norms of its constituents:

$$\|[y]\lambda\| = \sup_{x \in X} \|y\,(\lambda x)\|/\|x\| = \|y\| \sup_x |\lambda x|/\|x\| = \|y\|\|\lambda\|.$$

Consider the rank-one modification $Q := A - [y]\lambda$ of A. If $\lambda A^{-1}y = 1$, then $Q(A^{-1}y) = (A - [y]\lambda)A^{-1}y = y - y = 0$, i.e., Q is not invertible. Further, $\|A - Q\| = \|[y]\lambda\| = \|y\|\|\lambda\|$. Thus

$$\begin{aligned} d(A, \{Q \in bL(X,Y) : Q \text{ not invertible}\}) &\le \inf\{\|y\|\|\lambda\| : \lambda A^{-1}y = 1\} \\ &= 1/\sup_{\lambda,y} \frac{\lambda A^{-1}y}{\|\lambda\|\|y\|} \\ &= 1/\sup_{y} \frac{\|A^{-1}y\|}{\|y\|} \;=\; 1/\|A^{-1}\|, \end{aligned}$$

the second-last equality a consequence of the (IV.27)Hahn-Banach Theorem. H.P.(19) now finishes the proof. □

H.P.(20) Prove: *If* $\lim A_n = A$ (as elements of the nls $bL(X,Y)$), *and* A *is boundedly invertible, then* $\lim A_n^{-1} = A^{-1}$, i.e., A_n^{-1} exists for all sufficiently large n and $\|A_n^{-1} - A^{-1}\| \to 0$. (Hint: Prove first the useful identity $A^{-1} - Q^{-1} = A^{-1}(Q - A)Q^{-1}$.)