Analiza matematyczna II semestr wykład nr 5. (15.III 2021.)

5 Norma i iloczyn skalarny

Na tym etapie będziemy używali głównie normy euklidesowej $\|\cdot\|_2$ w przestrzeni wektorowej $\mathbb{R}^d$. Ale wkrótce zajmiemy się też normami w przestrzeniach funkcji, jak np. normą supremum $\|f\|_K := \sup\{|f(x)| : x \in K\}$ dla przestrzeni $C(K)$ funkcji ciągłych na zbiore zwartym $K \subset \mathbb{R}^d$, czy też normą całkową $\|f\|_1 := \int_a^b |f(t)|\,dt$, albo normą średniokwadratową $\|f\|_2 := \left(\int_a^b |f(t)|^2\right)^{\frac{1}{2}} dt$.

Warto więc poznać ogólną definicję:

Definicja. Przestrzenią unormowaną nazywamy parę $(X, \|\cdot\|)$, gdzie X jest przestrzenią wektorową (nad ciałem liczbowym $\mathbb{R}$ lub $\mathbb{C}$), zaś odwzorowanie $\|\cdot\| : X \ni v \mapsto \|v\| \in [0, +\infty)$, zwane normą, spełnia następujące postulaty:

- $\|v\| = 0 \Leftrightarrow v = 0$ —postulat tożsamości,
- $\|\alpha v\| = |\alpha|\|v\|$ —warunek jednorodności (z modułem)
- $\|v + w\| \leqslant \|v\| + \|w\|$ —nierówność trójkąta.

Mając normę określamy odległość dwu punktów v, w tej przestrzeni -jako liczbę $\|u-v\|$. Tak określona funkcja $d(u, v) := \|u-v\|$ nazywana jest metryką. Spełnia ona dla wszystkich wektorów tzw. aksjomaty przestrzeni metrycznej: warunek toższmości: $d(u, v) = 0 \Leftrightarrow u = v$, symetrii: $d(v, u) = d(u, v)$ i nierówność trójkąta: $d(u, v) \leqslant d(u, w) + d(w, v)$.[1]

Zbiór punktów w (oznaczany symbolem:)

$$K(u_0, R) = \{w \in X : d(u_0, w) < R\}$$

znajdujących się w odległości mniejszej od punktu $u_0 \in X$, niż liczba R, gdzie $R > 0$, nazywany jest **kulą** (otwartą) **o środku w** u_0**, o promieniu** R.

Średnicę zbioru niepustego $E \subset X$ definiujemy jako liczbę

$$\operatorname{diam}(E) := \sup\{d(u, w) : u, w \in E\}.$$

Niektóre normy są zdefiniowane przez **iloczyn skalarny.** W przestrzeni euklidesowej $\mathbb{R}^3$ dla wektorów $\vec{v}_1 = (x_1, y_1, z_1)$ oraz $\vec{v}_2 = (x_2, y_2, z_2)$ ich iloczyn skalarny zazwyczaj oznacza się symbolem $\vec{v}_1 \cdot \vec{v}_2$. Jest to liczba $x_1x_2 + y_1y_2 + z_1z_2$ i można wykazać, że jest ona równa $\|\vec{v}_1\|_2\|\vec{v}_2\|_2 \cos\angle(\vec{v}_1, \vec{v}_2)$, co jest bardzo pożyteczne w geometrii analitycznej (np. przy badaniu prostopadłości wektorów). Ale gdy wektorami będą funkcje, zapis $f_1 \cdot f_2$ jest zazwyczaj rozumiany jako funkcja $t \mapsto f_1(t)f_2(t)$. Głównie z tego powodu użyjemy innego oznaczenia iloczynu skalarnego- przy użyciu nawiasów graniastych:

$$\langle f, g\rangle.$$

Można też spotkać nawiasy okrągłe: $(f \,|\, g)$. Aby podać przykład iloczynu skalarnego funkcji, zdefiniujmy najpierw całkę Riemanna dla $f : [a, b] \to \mathbb{C}$. Każda taka funkcja o wartościach zespolonych może być zapisana w postaci $f(t) = a(t) + ib(t)$, gdzie $a = \Re f, b = \Im f$ oznaczają części: rzeczywistą i urojoną, zaś $i^2 = -1$. Mówimy, że taka funkcja jest całkowalna, gdy zarówno jej części rzeczywista, jak i urojona są całkowalne. W takim przypadku definiujemy

$$\int_a^b f(t)\,dt := \int_a^b \Re f(t)\,dt + i\int_a^b \Im f(t)\,dt.$$

Teraz dla f, g całkowalnych zdefiniujmy

$$\langle f, g\rangle := \int_a^b f(t)\overline{g(t)}\,dt. \tag{1.5}$$

[1] Faktycznie, $d(v, u) = \|-(u-v)\| = |-1|\|u-v\| = d(u, v)$ Natomiast $d(u, v) = \|u - w + w - v\| \leqslant \|u - w\| + \|w - v\|$.

Oczywiście, dla funkcji o wartościach rzeczywistych symbol sprzężenia jest opuszczany.

Na przykład, gdy $f(t) = \exp(int), g(t) = \exp(ikt)$, to

$$f(t)\overline{g(t)} = \exp(i(n-k)t) = \cos(nt - kt) + i\sin(nt - kt).$$

Jeśli całka liczona jest w granicach od $a = -\pi$ do $b = \pi$ - to tak określony iloczyn skalarny będzie równy zero gdy $n \neq k$ oraz równy 2π gdy $n = k$.

Obecność znaku sprzężenia dla funkcji g zapewnia nieujemność całki gdy $f = g$. Faktycznie, ponieważ dla $z \in \mathbb{C}$ mamy $z\bar{z} = |z|^2$, widzimy, że dla wcześniej zdefiniowanej normy mamy $\sqrt{\langle f, f\rangle} = \|f\|_2$, jeżeli f jest całkowalna. Aksjomatyka iloczynu skalarnego jest zawarta w następującej definicji: (sformułujemy ją w wariancie gdy ciałem skalarów są liczby zespolone i skomentujemy osobno-w przypadku rzeczywistym)

Definicja. Odwzorowanie $X \times X \ni (f, g) \to \langle f, g\rangle \in \mathbb{C}$ jest iloczynem skalarnym na zespolonej przestrzeni wektorowej X, gdy spełnione są warunki:

1. $\forall_{v\in X\setminus\{0\}}\ \langle v, v\rangle > 0$ (dodatnia określoność)
2. $\forall_{w_0\in X}$, odwzorowanie $X \ni v \mapsto \langle v, w_0\rangle \in \mathbb{C}$ jest liniowe,
3. $\langle w, v\rangle = \overline{\langle v, w\rangle}$ (skośna symetria)
   Odwzorowanie $X \ni w \mapsto \|w\| := \sqrt{\langle w, w\rangle}$ nazywamy normą tego iloczynu skalarnego.

W przypadku rzeczywistym warunki 2 oraz 3 oznaczają, że to odwzorowanie jest symetryczne i dwuliniowe. W przypadku zespolonym odwzorowania liniowe wzgl. pierwszej zmiennej i takie, że ich sprzężenie jest liniowe wzgl. drugiej zmiennej nazywamy półtoraliniowymi (ang. *sesquilinear maps*). Fakt, że $\|v\|$ faktycznie określa normę, zostanie za chwilę sprawdzony. Mamy więc wzory:

$$\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle \ \texttt{oraz} \ \ \langle v, \lambda w + \mu u\rangle = \bar{\lambda}\langle v, w\rangle + \bar{\mu}\langle v, u\rangle$$

Funkcja zmiennej $t \in \mathbb{R}$ określona przez iloczyn skalarny:

$$\langle u+tw, u+tw\rangle = \|u\|^2 + t\big(\langle u, w\rangle + \langle w, u\rangle\big) + t^2\|w\|^2 = t^2\|w\|^2 + 2t\Re\langle u, w\rangle + \|u\|^2 \tag{2.5}$$

jest stale nieujemna. Jest to trójmian kwadratowy zmiennej t, zatem jego wyróżnik Δ jest niedodatni: $\Delta \leqslant 0$. Ale tu $\Delta = (2\Re\langle u, w\rangle)^2 - 4\|u\|^2\|w\|^2$. Ten znak wyróżnika implikuje więc po spierwiastkowaniu stronami nierówność

$$|\Re\langle u, w\rangle| \leqslant \|u\|\|w\| \tag{3.5}$$

Wstawiając wartość $t = 1$ do (2.5) mamy

$$\|u + w\|^2 = \|u\|^2 + \|w\|^2 + 2\Re\langle u, w\rangle \leqslant \|u\|^2 + \|w\|^2 + 2\|u\|\|w\| = (\|u\| + \|w\|)^2,$$

czyli podniesioną stronami do kwadratu nierówność trójkąta.

Ponieważ $\langle \lambda u, \lambda u\rangle = \lambda\bar{\lambda}\|u\|^2$, pierwiastkując stronami otrzymamy warunek jednorodności dla normy. Postualt tożsamości też, rzecz jasna, zachodzi. Wykazaliśmy więc, że **iloczyn skalarny wyznacza normę.**

Jeśli w nierówności (3.5) zamiast w wstawimy wektor $e^{is}w$, gdzie $s \in \mathbb{R}$ jest tak dobrane, aby $\langle u, w\rangle = e^{is}|\langle u, w\rangle|$, to otrzymamy z (3.5) tak zwaną nierówność Schwarza. (W przypadku, gdy ciałem skalarów jest $\mathbb{R}$, (3.5) jest już tezą nierówności Schwarza). Nierówność ta jest niezmiernie przydatnym narzędziem w teorii przestrzeni Hilberta.

Twierdzenie. Dla iloczynu skalarnego i związanej z nim normy zachodzi następująca **nierówność Schwarza**

$$|\langle u, w\rangle| \leqslant \|u\|\|w\|. \tag{4.5}$$

Można nawet na podstawie znajomości normy odtworzyć iloczyn skalarny (poprzez tzw. wzór polaryzacyjny), ale nie każda norma pochodzi od jakiegoś iloczynu skalarnego. Można wykazać, że jest tak wtedy i tylko wtedy gdy zachodzi następująca "tożsamość równoległoboku":

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Na przykład w $\mathbb{R}^d$ mamy nieskończenie wiele różnych norm $\|w\|_p$ dla $p \in [1, +\infty)$ oraz dla $v = (x_1, x_2, \dots x_d)$ danych wzorem $\|v\|_p := \left(\sum_{k=1}^d |x_k|^p\right)^{\frac{1}{p}}$. Warunki: jednorodności i tożsamości są łatwe do sprawdzenia, nierówność trójkąta wykażemy później. Ale tylko dla jednej wartości p jest to norma pochodząca od iloczynu skalarnego, którym jest tu $\sum x_k y_k$. `Jakie jest to` p?

Całkowy iloczyn skalarny (1.5) definiuje jedną z norm całkowych (dla $p \geqslant 1$ taką normą z wykładnikiem p będzie $\|f\|_p := \left(\int_a^b |f(t)|^p\, dt\right)^{\frac{1}{p}}$. Wykazaliśmy, że jest to norma dla $p = 2$, zaś dla $p = 1$ to jest wniosek z całkowania stronami nierównosci $|f(t) + g(t)| \leqslant |f(t)| + |g(t)|$, ale dla innych p dowód jest nieco skomplikowany i przeprowadzimy go później. Wracając do normy euklidesowej w $\mathbb{R}^d$ -otrzymaliśmy dla niej nierówność Schwarza, którą w tym przypadku wcześniej wykazali Cauchy i Buniakowski, wygodnie będzie więc ją cytować jako (CBS):

$$\left|\sum_{k=1}^d x_k y_k\right| \leqslant \left(\sum_{k=1}^d |x_k|^2\right)^{\frac{1}{2}} \cdot \left(\sum_{k=1}^d |y_k|^2\right)^{\frac{1}{2}}. \qquad (CBS)$$

Norma (albo metryka) pozwalają na zdefiniowanie warunków:

- otwartości zbioru $D \subset X$: "Każdy punkt zbioru D zawiera się w zbiorze D wraz z pewną kulą"
- zbieżności ciągu wektorów $v_n \in X$ do wetora v -a mianowicie:

$$\lim_{n\to\infty} \|v_n - v_0\| = 0,$$

- ciągłości odwzorowania $F : D \subset \mathbb{R}^m \to X$ w punkcie $z \in D$:

$$\forall_{\epsilon>0} \exists_{\delta>0} \forall_{x\in D}\ \|x - z\| < \delta \Rightarrow \|F(x) - F(z)\| < \epsilon,$$

- warunków: Lipschitza i jednostajnej ciągłości takiego odwzorowania F

Norma jest odwzorowaniem ciągłym -bo spełnia warunek Lipschitza ze stałą 1. Taki warunek Lipschitza, to (będąca prostym wnioskiem z pierwszej) **Druga Nierówność Trójkąta:**

$$\big|\|x\| - \|y\|\big| \leqslant \|x - y\|.$$

Dowód. Faktycznie, jest to nierówność typu $|a| \leqslant M$, co jest równoważne koniunkcji dwu nierówności: $-M \leqslant a \leqslant M$. Każdą z tych dwu nierówności w naszym przypadku sprowadzamy do zwykłej nierówności trójkąta. Na przykład, $-\|x - y\| \leqslant \|x\| - \|y\| \Leftrightarrow \|y\| \leqslant \|x\| + \|y - x\|$, a ostatnia nierówność, to nierówność trójkąta $\|u + w\| \leqslant \|u\| + \|w\|$ dla $u = x, w = y - x$. □

W zasadzie mamy tu pełną analogię z odpowiednimi definicjami w przypadku jednowymiarowym, gdzie zamiast normy występował moduł. Podobnie można też definiować granicę funkcji w punkcie skupienia dziedziny i pochodną $\gamma'(t)$ dla krzywej $\gamma : [a, b] \to \mathbb{R}^d$. Jak łatwo sprawdzić, będzie to zestawienie pochodnych poszczególnych współrzędnych. Już nawet dla $d = 2$ mamy jednak poważną różnicę:

Ani twierdzeń Rolle'a, czy Lagrange'a, ani twierdzenia Cauchy'ego o wartości średniej nie da się przenieść na przypadek funkcji o wartościach wektorowych. Przykładem jest $\gamma(t) = (\cos t, \sin t) \in \mathbb{R}^2$ dla $t \in [0, 2\pi]$. Po pełnym obiegu okręgu wracamy do punktu wyjścia, ale pochodna w żadnym punkcie nie jest tu równa zero, bo $\|\gamma'(t)\|_2 = \|(-\sin t, \cos t)\|_2 = 1$ dzięki jedynce trygonometrycznej. Jednak ważny wniosek z twierdzenia Lagrange'a zwany **twierdzeniem o przyrostach** ma swój wielowymiarowy odpowiednik. Sformułujemy go dla normy euklidesowej w $\mathbb{R}^d$ (chociaż można je wykazać dla dowolnej przestrzeni unormowanej).

Twierdzenie o Przyrostach. Gdy odwzorowanie $\gamma : [a, b] \to \mathbb{R}^d$ jest ciągłe w $[a, b]$ i różniczkowalne w przedziale otwartym (a, b), to

$$\|\gamma(b) - \gamma(a)\|_2 \leqslant (b - a) \sup_{t\in(a,b)} \|\gamma'(t)\|_2. \tag{5.5}$$

Dowód. Niech $r := \|\gamma(b) - \gamma(a)\|$. Jeśli $r = 0$, to teza (5.5) jest oczywista. W przeciwnym przypadku niech $v = \frac{1}{r}(\gamma(b) - \gamma(a))$. Dzięki jednorodnosci normy, mamy $\|v\|_2 = \frac{1}{r}\|\gamma(b) - \gamma(a)\|_2 = 1$. Jeśli $\gamma(t) = (x_1(t), \ldots, x_n(t))$, zaś $v = (v_1, \ldots, v_d)$ jest ustalonym wektorem, to funkcja

$$[a, b] \ni t \mapsto \Phi(t) := \langle\gamma(t), v\rangle$$

spełnia założenia tw. Lagrange'a o wartości średniej: jest ciągła w $[a, b]$ i dla każdego $t \in (a, b)$ ma pochodną. Faktycznie, $\Phi'(t)$, jako pochodna z $\Phi(t) = \sum_{k=1}^{d} x_k(t)v_k$ istnieje i jest równa iloczynowi skalarnemu

$$\Phi'(t) = \langle\gamma'(t), v\rangle = \sum_{k=1}^{d} x_k'(t)v_k. \tag{6.5}$$

Z twierdzenia Lagrange'a dla funkcji (o wartościach skalarnych) Φ, istnieje taki punkt $t_0 \in (a, b)$, w którym $\frac{\Phi(b)-\Phi(a)}{b-a} = \Phi'(t_0)$. Prawa strona tej równości jest nie większa, niż

$$\sup_{t\in(a,b)} |\Phi'(t)| \leqslant \sup_{t\in(a,b)} \|\gamma'(t)\|_2\|v\|_2.$$

Ostatnia z nierówności, to nierówność Schwarza zastosowana do wzoru (6.5). Wiec uwzględniając równość $\|v\|_2 = 1$ mamy

$$\Phi(b) - \Phi(a) \leqslant (b - a) \sup_{t\in[a,b]} \|\gamma'(t)\|. \tag{7.5}$$

Ale dzięki liniowości iloczynu skalarnego wzgl. pierwszej zmiennej, $\Phi(b)-\Phi(a) = \langle\gamma(b)-\gamma(a), v\rangle$. Teraz przywołując definicję wektora v i wyłączając $\frac{1}{r}$ przed iloczyn skalarny widzimy, że $\langle\gamma(b) - \gamma(a), v\rangle = \frac{1}{r}\|\gamma(b) - \gamma(a)\|_2^2 = r$. Wstawiając otrzymane równości do lewej strony nierówności (7.5) dostajemy tezę. $\square$

Załóżmy teraz, że $\gamma : [a, b] \to \mathbb{R}^d$ jest krzywą klasy C^1. Gdy $a \leqslant s \leqslant b$, oznaczmy przez $L(s)$ długość fragmentu tej krzywej dla $t \leqslant s$, czyli restrykcji (= zawężenia) krzywej γ do odcinka $[a, s]$. Oczywiście, długością $\ell(\gamma)$ całej naszej krzywej jest liczba $L(b) = \int_a^b L'(s)\, ds$, a ostatnia równość wynika ze wzoru Newtona - Leibniza. Dla dowodu podanego na poprzednim wykładzie wzoru $\ell(\gamma) = \int_a^b \|\gamma'(t)\|_2\, dt$ wystarczy więc wykazać następujący lemat:

Lemat o różniczce długości krzywej. Pochodna z funkcji $L(t)$ dana jest wzorem $L'(t) = \|\gamma'(t)\|_2$, czyli $d\ell = \|\gamma'(t)\|_2\, dt$.

Dowód. wykorzystamy (analogiczny do addytywnej zależności całki od drogi całkowania) fakt, że dla $h > 0$ różnica $L(t + h) - L(t)$ jest długością krzywej $\gamma|_{[t,t+h]}$ powstałęj przez restrykcję γ do przedziału $[t, t+h]$. Jeśli mamy łamaną $\mathcal{L}_m$ wpisaną w krzywą $\gamma|_{[t,t+h]}$ i $\mathcal{L}_m$ odpowiada punktom podziału $t_0 = t < t_1 < \cdots < t_m = t + h$, to ma ona długość $\sum_{j=1}^{m} \|\gamma(t_j) - \gamma(t_{j-1})\|_2$ =suma norm przyrostów γ na kolejnych odcinkach tego podziału. Każdy z przyrostów szacujemy dzięki poprzedniemu twierdzeniu przez $(t_j-t_{j-1}) \sup_{s\in[t,t+h]} \|\gamma'(t)\|$, a ich sumę -przez $h \cdot \sup_{s\in[t,t+h]} \|\gamma'(t)\|$. Iloraz różnicowy $\frac{1}{h}(L(t + h) - L(t))$ jest więc nie większy, niż to supremum, które przy $h \to 0^+$ zmierza do $\|\gamma'(t)\|_2$, dzięki ciągłości γ' i normy. Ale mamy też oszacowanie ilorazu $\frac{1}{h}(L(t+h)-L(t))$ z dołu -przez $\frac{1}{h}\|\gamma(x+h)-\gamma(x)\|$, bo długość tego kawałka krzywej jest większa lub równa od długości cięciwy -czyli odcinka o końcach $\gamma(t), \gamma(t + h)$. Czyli $\frac{1}{h}(L(t+h) - L(t)) \geqslant \|\frac{1}{h}(\gamma(t + h) - \gamma(t))\|_2$. Dzięki ciągłości normy, to zmierza do $\|\gamma'(t)\|$ gdy $h \to 0^+$. Dla $h \to 0^-$ -jak w dowodzie wzoru Newtona-Leibniza sprowadzamy to do poprzedniego przypadku (gdzie $h > 0$).$\square$