Następne cztery wykłady poświęcone będą badaniu dodatkowych możliwości jakie daje wprowadzenie na przestrzeni wektorowej na której pracujemy iloczynu skalarnego. Najpierw ogólne uwagi o dodatkowych strukturach:

Z punktu widzenia fizyka, teoretyka i doświadczalnika, matematyka to pewne narzędzie, które wykorzystujemy w naszej pracy. Inaczej jej używają teoretycy, a inaczej doświadczalnicy. Dla teoretyka matematyka jest językiem dzięki któremu opisujemy rzeczywistość fizyczną. Tworzymy wyidealizowane modele pewnych zjawisk starając się te zjawiska zrozumieć poprzez te modele. Język musi więc być jak najlepiej dopasowany do sytuacji. Jeśli język będzie zbyt bogaty, model będzie miał własności, których nie ma rzeczywistość. Z drugiej strony jeśli model jest zbyt ubogi to nie opiszemy i nie zrozumiemy być może istotnych szczegółów. Kolejna sprawa to poziom trudności: naturalne jest, że wolelibyśmy żeby rachunki, które robimy w ramach danego modelu były w miarę proste. Są przykłady teorii w których zmiana języka znacząco uprościła techniki rachunkowe.

Koniec indoktrynacji: przechodzimy do meritum.

DEFINICJA: Iloczynem skalarnym na przestrzeni wektorowej V nad ciałem $\mathbb{K}$ nazywamy odwzorowanie

$(\cdot|\cdot) : V \times V \longrightarrow \mathbb{K}$ mające następujące własności

(1) $\forall v \in V, u_1, u_2 \in V, \lambda_1, \lambda_2 \in \mathbb{K}$ zachodzi

$$(v, \lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 (v|u_1) + \lambda_2 (v, u_2)$$

liniowość ze względu na drugi argument

(2) $\mathbb{K} = \mathbb{R}$

$$\forall u, v \in V \quad (u|v) = (v|u)$$

symetria

$\mathbb{K} = \mathbb{C}$

$$(u, v) = \overline{(v, u)}$$

hermitowskość

Charles Hermite

1822 – 1901

Francja

(3) $\forall u \in V \quad u \neq 0 \qquad (u|u) > 0$

Punkt (3) w szczególności oznacza dla $\mathbb{K} = \mathbb{C}$, że $(u|u)$ musi być rzeczywiste. Ponadto, także dla $\mathbb{K} = \mathbb{C}$ mamy

$$(\lambda u|v) = \overline{(v|\lambda u)} = \overline{\lambda (v|u)} = \bar{\lambda}\, \overline{(v|u)} = \bar{\lambda} (u|v)$$

ogólniej

$$(\lambda_1 u_1 + \lambda_2 u_2 | v) = \bar{\lambda}_1 (u_1|v) + \bar{\lambda}_2 (u_2|v)$$

Własność tę nazywa się antyliniowością ze względu na pierwszy argument.

Widać wobec tego, że w przestrzeni rzeczywistej iloczyn skalarny jest formą dwuliniową symetryczną, dodatnio określoną (czyli o sygnaturze $(n,0)$) w szczególności także niezdegenerowaną. W przestrzeni zespolonej jest inaczej: forma jest półtoraliniowa, to znaczy antyliniowa ze względu na pierwszy i liniowa ze względu na drugi argument. Zamiast symetrii mamy warunek hermitowskości. Warunek (3) implikuje, że forma jest niezdegenerowana, co w tym przypadku oznacza np., że odwzorowanie

$V \ni v \longmapsto (v|\cdot) \in V^*$ jest antyliniowym izomorfizmem.

PODSTAWOWE PRZYKŁADY:

$\mathbb{R}^n: \ (x|y) = \sum_i x^i y^i \qquad \mathbb{C}^n: \ (x|y) = \sum_i \overline{x^i} y^i$

$\mathbb{R}_n[\cdot]: \ (u|v) = \int_a^b u(t) v(t)\, dt \qquad \mathbb{C}_n[\cdot]: \ (u|v) = \int_a^b \bar{u}(t) v(t)\, dt$

$M_{n\times n}(\mathbb{R}): \ (X|Y) = tr(X^T Y) \qquad M_{n\times n}(\mathbb{C}): \ (X|Y) = tr(X^\dagger Y)$

Sprzężenie hermitowskie macierzy to transpozycja + sprzężenie zespolone wyrazów.

Podstawową konsekwencją wprowadzenia iloczynu skalarnego w przestrzeni jest możliwość mierzenia długości wektorów i obliczania odległości między wektorami. Nasza przestrzeń wektorowa staje się przestrzenią unormowaną z odpowiednią metryką pochodzącą od normy.

DEFINICJA: $x \in V$ $\quad \|x\| = \sqrt{(x|x)}$ $\quad \|x\|$ nazywamy długością albo normą wektora x.

Własności tej normy są takie jak trzeba, tzn $\|x\| = 0 \Leftrightarrow x = 0$
niezdegenerowanie

$\|\lambda x\| = |\lambda| \, \|x\|$ (dodatnia jednorodność)

FAKT: Nierówność Schwarza: $|(x|y)| \leqslant \|x\| \cdot \|y\|$

DOWÓD: Załóżmy, że x i y są różne od 0. Jeśli któryś jest zero nierówność i tak jest prawdziwa, bo po obu stronach są zera. Rozważmy

$(tx - y | tx - y)$ wiadomo, że to jest $\geqslant 0$ dla dowolnego t

$$(tx-y|tx-y) = t^2(x|x) - 2t(x|y) + (y|y) = t^2\|x\|^2 - 2t(x|y) + \|y\|^2 \geqslant 0$$

nad $\mathbb{R}$ $\qquad$ dla dowolnego t

$$\Delta = 4(x|y)^2 - 4\|x\|^2\|y\|^2 \leqslant 0$$

$$(x|y)^2 \leqslant \|x\|^2\|y\|^2$$

$e^{-i\varphi}(x|y) = |(x|y)|$

$(e^{i\varphi}x|y) = |(x|y)|$

Oznacza to, że istotnie $|(x|y)| \leqslant \|x\| \|y\|$

Dla $\mathbb{K} = \mathbb{C}$ trzeba rozpatrzeć inną funkcję. Weźmy φ: $(x|y) = e^{i\varphi}|(x|y)|$ i rozważmy

$$0 \geqslant f(t) = (te^{i\varphi}x - y | te^{i\varphi}x - y) = t^2\|x\|^2 + \|y\|^2 - \underbrace{t(e^{i\varphi}x|y) - t(y|e^{i\varphi}x)}_{2t\,\mathrm{Re}(e^{i\varphi}x|y) = 2t|(x|y)|}$$ i dalej j.w.

Konsekwencje nierówności Schwarza są dwie dla przestrzeni rzeczywistych i jedna dla przestrzeni zespolonych:

$K = \mathbb{R}$

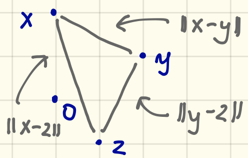
$d(x,y)=\|x-y\|$ jest odległością. Jak zwykle, to co trzeba udowodnić to nierówność trójkąta

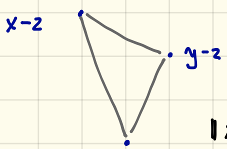
$$d(x,y) \leq d(x,z)+d(y,z)$$

$$\|x-y\| \leq \|x-z\| + \|y-z\|$$

Zauważmy najpierw, że wystarczy przyjąć $z=0$, bo przesunięcie zachowuje d:

$$d(x-z,y-z) = \|x-z-y+z\| = \|x-y\| = d(x,y)$$





$$x' = x-z$$
$$y' = y-z$$

$$\|x'-y'\| = \|x-y\|$$
$$\|x'\| = \|x-z\|$$
$$\|y'\| = \|y-z\|$$

$$\|x-y\| \leq \|x-z\| + \|y-z\|$$
$$\|x'-y'\| = \|x'\| + \|y'\|$$

Opuszczamy primy...

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

$$-\|x\|\,\|y\| \leq (x|y) \leq \|x\|\,\|y\| \quad /\cdot(-2)$$

$$-2\|x\|\,\|y\| \leq -2(x|y) \leq 2\|x\|\,\|y\|$$

$$\|x\|^2+\|y\|^2-2\|x\|\,\|y\| \leq \|x\|^2+\|y\|^2-2(x|y) \leq \|x\|^2+\|y\|^2+2\|x\|\,\|y\|$$

$$(\|x\|-\|y\|)^2 \leq \|x-y\|^2 \leq (\|x\|+\|y\|)^2$$

$$\big|\,\|x\|-\|y\|\,\big| \leq \underbrace{\|x-y\| \leq \|x\|+\|y\|}_{\text{nierówność trójkąta}}$$

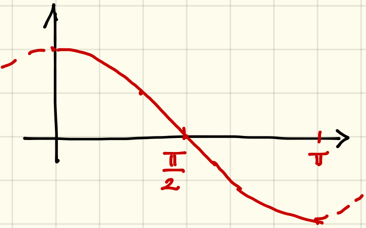
Konsekwencja nr 2 to możliwość zdefiniowania kąta między wektorami:
skoro

$$|(x|y)| \leqslant \|x\|\,\|y\| \quad \text{to} \quad -\|x\|\cdot\|y\| \leqslant (x|y) \leqslant \|x\|\,\|y\|$$

więc

$$-1 \leqslant \frac{(x|y)}{\|x\|\,\|y\|} \leqslant 1$$

$$= \cos\alpha \qquad \alpha \in [0,\pi]$$



W szczególności wiadomo kiedy wektory są prostopadłe: $x \perp y \iff (x|y)=0$

Dla $\mathbb{K}=\mathbb{C}$ kwestia odległości jest bez zmian, tzn przestrzeń zespolona z hermitowskim iloczynem skalarnym jest metryczna. $(x|y)$ ma wartości zespolone, wobec tego nie definiujemy kąta między wektorami.
Nadal jednak mamy pojęcie prostopadłości.

Dopełnienie ortogonalne $(\mathbb{K}=\mathbb{R},\mathbb{C})$

Niech V będzie przestrzenią wektorową a $\mathcal{U}\subset V$ podprzestrzenią w V.
Definiujemy

$$\mathcal{U}^{\perp} = \{v\in V : \forall_{u\in\mathcal{U}}\ (u|v)=0\}$$

FAKT: (1) $\mathcal{U}^{\perp}$ jest podprzestrzenią wektorową

(2) $V=\mathcal{U}\oplus\mathcal{U}^{\perp}$

Dowód: (1) niech $v, v'\in\mathcal{U}^{\perp}$ wtedy dla $\lambda v+\mu v'$ mamy

$$(u|\lambda v+\mu v') = \lambda(u|v)+\mu(u|v') = 0+0=0.$$

(2) Zacznijmy od $\mathcal{U} \cap \mathcal{U}^{\perp} = \{0\}$, istotnie jeśli $u \in \mathcal{U}$ i $u \in \mathcal{U}^{\perp}$ to $(u|u)=0$, zatem $u=0$, bo mamy niezdegenerowany il. skalarny.

Policzmy teraz wymiar $\mathcal{U}^{\perp}$. Skorzystamy z odwzorowania

$v \xmapsto{G} (\ |v)$ $V \longrightarrow V^*$. Jest to liniowy izomorfizm. Zauważmy, że jeśli $v \in \mathcal{U}^{\perp}$ to $G(v) \in \mathcal{U}^{0}$ i odwrotnie, jeśli $\alpha \in \mathcal{U}^0$ to $G^{-1}(\alpha) \in \mathcal{U}^{\perp}$ ten $\mathcal{U}^{\perp} = G^{-1}(\mathcal{U}^0)$. Wiadomo, że $\dim \mathcal{U}^0 = \dim V - \dim \mathcal{U}$. Z izomorfizmu $\dim \mathcal{U}^{\perp} = \dim \mathcal{U}^0$. Zatem

$$\dim \mathcal{U} + \dim \mathcal{U}^{\perp} = \dim \mathcal{U} + \dim V - \dim \mathcal{U} = \dim V$$ oznacza to, że

$$\mathcal{U} + \mathcal{U}^{\perp} = V. \quad \blacksquare$$

Odległości między zbiorami: W przestrzeni metrycznej zdefiniowana jest odległość między punktem a zbiorem oraz między zbiorami

$$(M,d) \quad x \in M,\ A \subset M \quad d(x,A) = \inf\{d(x,y) : y \in A\}$$

$$A,B \subset M \qquad d(A,B) = \inf\{d(x,y) : x \in A,\ y \in B\}$$

Zobaczmy jak to się manifestuje w przestrzeniach z iloczynem skalarnym dla szczególnych podzbiorów.

DEFINICJA Niech V będzie przestrzenią wektorową, $\mathcal{U}$ podprzestrzenią. Podprzestrzenią afiniczną modelowaną na $\mathcal{U}$ nazywamy podzbiór $\mathcal{A} \subset V$ taki, że istnieje $a \in \mathcal{A}$

$$\mathcal{A} = a + \mathcal{U} = \{a+u \mid u \in \mathcal{U}\}$$

Oznacza to między innymi, że dla $a,b \in \mathcal{A}$ $a-b \in \mathcal{U}$. Ponadto jeśli mamy jedno takie a, to każde a' postaci $a'+u$, $u \in \mathcal{U}$ jest równie dobre.

Przykład: Niech $U \subset V$: $\dim U = 1$ $L = l + U$ nazywamy prostą afiniczną w V. Prosta ta przechodzi przez l

Znajdźmy wzór na odległość punktu od prostej: (analitycznie)

$$L = l + U \quad U = \langle u \rangle \quad x \in V \qquad d(x, L) = ?$$

Z definicji $d(x,L) = \inf\{d(x,y),\ y \in L\}$ $y = l + tu$ dla $t \in \mathbb{R}$

$d(x, l+tu) = \|x - l - tu\| = (x-l-tu \mid x-l-tu)^{1/2}$ Odległość jest dodatnia - możemy szukać minimum kwadratu:

$f(t) = (x-l-tu \mid x-l-tu)$ funkcja ciągła i różniczkowalna na $\mathbb{R}$ (kwadratowa). Szukamy minimum:

$$f(t) = (x-l \mid x-l) - 2(x-l \mid u)t + t^2(u \mid u) = \|x-l\|^2 - 2t(x-l \mid u) + t^2\|u\|^2$$

$$f'(t) = -2(x-l \mid u) + 2t\|u\|^2 = 0 \qquad t = \frac{(x-l \mid u)}{\|u\|^2}$$

$$d(x,L) = d\Big(x, \underbrace{l + \frac{(x-l \mid u)}{\|u\|^2}u}_{\text{co to jest?}}\Big)$$

$$y_0 = l + \frac{(x-l \mid u)}{\|u\|^2}u = l + \Big(x-l \,\Big|\, \underbrace{\frac{u}{\|u\|}}_{e}\Big)\frac{u}{\|u\|} = l + (x-l \mid e)e$$

$e \leftarrow$ jednostkowy kierunkowy wektor U.

Co jest szczególnego w y_0? Obserwacja

$$(x-y_0 \mid e) = (x - l - (x-l \mid e)e \mid e) = (x-l \mid e) - (x-l \mid e)\underbrace{(e \mid e)}_{1} = 0.$$

ten $x-y_0 \in U^\perp$. Zamiast szukać minimum funkcji rzeczywistej możemy szukać y_0 takiego, że $y_0 \in L$ i $x-y_0 \in U^\perp$

$$y_0 = l + ut \qquad (x-y_0 \mid u) = 0$$

$$(x-l-ut \mid u) = 0 \rightarrow \text{ten sam wzór na } y_0.$$

Odległość między prostymi: (Można analitycznie, ale lepiej pomyśleć geometrycznie)

$$L = l+tu, \ P = p+sv \qquad d(L,P) = \inf d(l+tu, p+sv) = \inf \|(l-p)+(tu-sv)\|$$

pewien punkt

podprzestrzeń wektorowa rozpięta przez $\langle u,v\rangle$

Musimy więc umieć policzyć odległość punktu od podprzestrzeni wektorowej:

x_0 - punkt $\quad U$ - podprzestrzeń

Mając podprzestrzeń mamy $V = U \oplus U^\perp$ i mamy rzut na U wzdłuż $U^\perp$ (rzut ortogonalny!)

$$d(x_0,U)^2 = \inf_{y\in U} \|x_0, y\|^2 = \inf \|\underbrace{x_0 - P_U x_0}_{} + \underbrace{P_U x_0 - y}_{}\|^2 = \inf\left(\|x_0 - P_U x_0\|^2 + \underset{=0}{\|P_U x_0 - y\|}\right)$$

są prostopadłe!

$$= \|x_0 - P_U x_0\| = \|P_{U^\perp} x_0\|$$

Wracamy do odległości prostych:

$$d(L,P) = \|l-p - P_{\langle u,v\rangle}(l-p)\| = \|P_{\langle u,v\rangle^\perp}(l-p)\|$$

Ogólnie odległość p.p. afinicznych:

$$\mathcal{A}_1 = x_1 + \mathcal{U}_1 \quad \mathcal{A}_2 = x_2 + \mathcal{U}_2 \qquad d(\mathcal{A}_1, \mathcal{A}_2) = \| (x_1 - x_2) - P_{\mathcal{U}_1 + \mathcal{U}_2}(x_1 - x_2) \| =$$
$$= \| P_{(\mathcal{U}_1 + \mathcal{U}_2)^\perp}(x_1 - x_2) \|$$

Struktury jakimi posługiwaliśmy się do tej pory są nam bliskie, ponieważ używamy ich na codzień. Uczymy się już w szkole podstawowej o kątach między prostymi. Teraz jedynie upewniamy się, że wszystko to związane jest z iloczynem skalarnym i wobec tego może być uogólnione na bardziej wyrafinowane przestrzenie niż $\mathbb{R}^n$.

Sprawdźmy teraz jak wyglądałby świat, gdyby obowiązywała w nim całkiem inna geometria. Dla ułatwienia będziemy pracować w przestrzeni trójwymiarowej (zamiast 4). Mamy

$V : \dim V = 3$ $\quad \eta$ - forma dwuliniowa, niezdegenerowana o sygnaturze $(+--)$. Oczywiście całej struktury nie przeanalizujemy z braku czasu ale przyjrzyjmy się przynajmniej pewnym elementom:

(1) Odwzorowanie $\quad N : V \longrightarrow V^* \quad v \longmapsto \eta(v, \cdot)$ jest izomorfizmem, bo η jest niezdegenerowana. Istnieje wobec tego pojęcie $\mathcal{U}^\perp$ dla podprzestrzeni, tzn

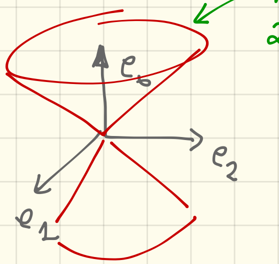
$$\mathcal{U}^\perp = N^{-1}(\mathcal{U}^\circ)$$ Nie jest jednak prawdą, że $V = \cancel{\mathcal{U} \oplus \mathcal{U}^\perp}$

ponieważ istnieją wektory tzw. zerowe, tzn $v \in V : \; v \neq 0$ i $\eta(v, v) = 0$. wektor może więc być prostopadły do samego siebie. Co to są za wektory? Weźmy dowolną bazę diagonalizującą η. Dla wygody zadbajmy aby wektory bazy były unormowane, tzn aby współczynniki diagonalne były ± 1.

$$(e_0, e_1, e_2) \quad [\eta]_e = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix}$$

Weźmy $v = x^0 e_0 + x^1 e_1 + x^2 e_2$. Wtedy $\eta(v,v) = (x^0)^2 - (x^1)^2 - (x^2)^2$. Warunek $\eta(v,v)=0$ oznacza

$$x^0 = \pm\sqrt{(x^1)^2 + (x^2)^2} \quad \text{— stożek}$$



Istnieje nawet baza składająca się wyłącznie z wektorów zerowych.

Wektory we wnętrzu stożka mają własność $\eta(v,v)>0$ i nazywają się czasowe, wektory na zewnątrz nazywają się przestrzenne i mają własność $\eta(v,v)<0$.

Wybierzmy jakiś wektor czasowy, np e_0, wtedy $\langle e_0\rangle^\perp = \langle e_1, e_2\rangle$ czyli w miarę normalnie. η indukuje na $\langle e_1, e_2\rangle$ iloczyn skalarny $g = -\eta|_{\langle e_1, e_2\rangle}$. Ma sygnaturę $(++)$ i jest zwykłym euklidesowym iloczynem skalarnym. Wybierzmy inny wektor czasowy, np

$\tilde{f}_0 = e_0 + \frac{1}{2} e_1$ Wtedy też istnieje $\langle f_0\rangle^\perp$ takie, że $V = \langle f_0\rangle \oplus \langle f_0\rangle^\perp$

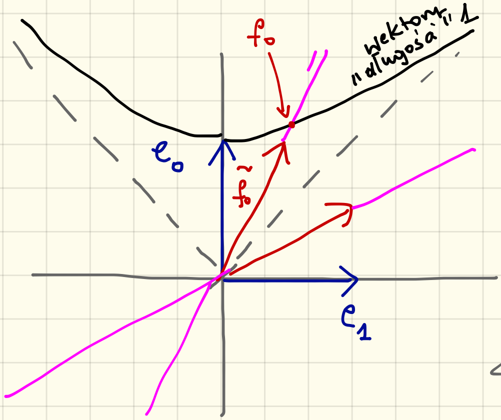
$\langle f_0\rangle^\perp = \langle e_2, \tilde{f}_1 = \frac{1}{2} e_1 + e_0\rangle$

Unormowane:

$$f_0 = \frac{2\sqrt{3}}{3} e_0 + \frac{\sqrt{3}}{3} e_1$$

$$f_1 = \frac{\sqrt{3}}{3} e_1 + \frac{2\sqrt{3}}{3} e_0$$

$$f_2 = e_2$$



$\langle f_1, f_2\rangle$