2 - 9 - REELLE INDREPRODUKTROM - LF

$\boxed{1}$ La oss begynne med skalarproduktet. For det første er

$$\mathbf{x}^T\mathbf{y} = \sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} y_k x_k = \mathbf{y}^T\mathbf{x},$$

for det andre er

$$\mathbf{x}^T\mathbf{x} = \sum_{k=1}^{n} x_k x_k = \sum_{k=1}^{n} x_k^2 \geq 0$$

og kun lik null for nullvektoren, og for det tredje er

$$(a\mathbf{x} + b\mathbf{y})^T\mathbf{z} = a\mathbf{x}^T\mathbf{z} + b\mathbf{y}^T\mathbf{z}.$$

Så til det andre indreproduktet. For det første er

$$\int_a^b x(t)y(t)\ dt = \int_a^b y(t)x(t)\ dt =$$

for det andre er

$$\int_a^b x(t)x(t)\ dt = \int_a^b x^2(t)\ dt \geq 0$$

siden x^2 alltid er positiv, og for det tredje er

$$\int_c^d (ax(t)+by(t))z(t)\ dt = \int_c^d ax(t)z(t)+by(t)z(t)\ dt = a\int_c^d x(t)z(t)\ dt + b\int_c^d y(t)z(t)\ dt.$$

$\boxed{2}$ **1:** Vi bruker definisjonen på lengde og første aksiom, og beregner

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 =&(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})\\ =&(\mathbf{x},\mathbf{x}) + (\mathbf{x},\mathbf{y}) + (\mathbf{y},\mathbf{x}) + (\mathbf{y},\mathbf{y})\\ =&(\mathbf{x},\mathbf{x}) + 2(\mathbf{x},\mathbf{y}) + (\mathbf{y},\mathbf{y})\\ =&\|\mathbf{x}\|^2 + 2(\mathbf{x},\mathbf{y}) + \|\mathbf{y}\|^2.\end{aligned}$$

På den ene side: Dersom

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

kan vi trekke denne fra likningen over og se at

$$2(\mathbf{x},\mathbf{y}) = 0,$$

som betyr at $\mathbf{x}$ og $\mathbf{y}$ er ortogonale. På den andre: Dersom $(\mathbf{x},\mathbf{y}) = 0$, får vi

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

2: Cauchy-Schwarz' ulikhet: $\quad |\,(\mathbf{x},\mathbf{y})\,| \leq \|\mathbf{x}\|\|\mathbf{y}\|$

La t være en skalar. Vi beregner først

$$\|\mathbf{x}+t\mathbf{y}\|^2 = (\mathbf{x},\mathbf{x}) + 2t(\mathbf{x},\mathbf{y}) + t^2(\mathbf{y},\mathbf{y}).$$

Denne kvadratiske funksjonen av t er alltid positiv, og dersom $\mathbf{y}$ ikke er nullvektoren, har den et nullpunkt i

$$t = -\frac{(\mathbf{x},\mathbf{y})}{(\mathbf{y},\mathbf{y})},$$

som gir

$$0 \leq (\mathbf{x},\mathbf{x}) - 2\frac{(\mathbf{x},\mathbf{y})^2}{(\mathbf{y},\mathbf{y})} + \left(\frac{(\mathbf{x},\mathbf{y})}{(\mathbf{y},\mathbf{y})}\right)^2(\mathbf{y},\mathbf{y}) = (\mathbf{x},\mathbf{x}) - \frac{(\mathbf{x},\mathbf{y})^2}{(\mathbf{y},\mathbf{y})}$$

eller

$$(\mathbf{x},\mathbf{y})^2 \leq (\mathbf{x},\mathbf{x})(\mathbf{y},\mathbf{y})$$

eller

$$|(\mathbf{x},\mathbf{y})| \leq \|\mathbf{x}\|\|\mathbf{y}\|.$$

3: Vi beregner

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 =&(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})\\ =&\|\mathbf{x}\|^2 + 2(\mathbf{x},\mathbf{y}) + \|\mathbf{y}\|^2\\ \leq&\|\mathbf{x}\|^2 + 2|(\mathbf{x},\mathbf{y})| + \|\mathbf{y}\|^2\\ \leq&\|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\|+\|\mathbf{y}\|)^2\end{aligned}$$

slik at

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

4: Vi antar at

$$\mathbf{x} = \sum_{k=1}^{n} = c_k\mathbf{v}_k$$

Vi tar indreproduktet med $\mathbf{v}_m$ og bruker ortogonaliteten, og får

$$(\mathbf{x},\mathbf{v}_m) = \left(\sum_{k=1}^{n} c_k\mathbf{v}_k, \mathbf{v}_m\right) = \sum_{k=1}^{n} c_k(\mathbf{v}_k,\mathbf{v}_m) = c_m(\mathbf{v}_m,\mathbf{v}_m)$$

slik at

$$c_m = \frac{(\mathbf{x},\mathbf{v}_m)}{(\mathbf{v}_m,\mathbf{v}_m)}$$

for alle m. Dette betyr at

$$\mathbf{x} = \sum_k \frac{(\mathbf{x},\mathbf{v}_k)}{(\mathbf{v}_k,\mathbf{v}_k)}\mathbf{v}_k$$

5: Først kan vi se at $\mathbf{x}-\mathbf{z}$ står ortogonalt på V, siden

$$\begin{aligned}(\mathbf{x}-\mathbf{z},\mathbf{v}_m) &= (\mathbf{x},\mathbf{v}_m) - (\mathbf{z},\mathbf{v}_m)\\ &= (\mathbf{x},\mathbf{v}_m) - \left(\sum_k \frac{(\mathbf{x},\mathbf{v}_k)}{(\mathbf{v}_k,\mathbf{v}_k)}\mathbf{v}_k, \mathbf{v}_m\right)\\ &= (\mathbf{x},\mathbf{v}_m) - (\mathbf{x},\mathbf{v}_m) = 0\end{aligned}$$

for alle $\mathbf{v}_m$. Dersom $\mathbf{y}$ er en annen vektor i V, ligger også $\mathbf{y}-\mathbf{z}$ i V, og da står $\mathbf{y}-\mathbf{z}$ og $\mathbf{x}-\mathbf{z}$ ortogonalt på hverande. Pytagoras' setning med retningen på $\mathbf{y}$ byttet om gir

$$\begin{aligned}\|\mathbf{x}-\mathbf{y}\|^2 &= \|\mathbf{x}-\mathbf{z}-(\mathbf{y}-\mathbf{z})\|^2 \\ &= \|\mathbf{x}-\mathbf{z}\|^2 + \|\mathbf{y}-\mathbf{z}\|^2 \geq \|\mathbf{x}-\mathbf{z}\|^2.\end{aligned}$$

for alle $\mathbf{y} \in V$.

$\boxed{3}$ Her gjelder det å huske trigonometriske identiteter. En klassiker er

$$\cos(\alpha \mp \beta) = \cos(\alpha)\cos(\beta) \pm \sin(\alpha)\sin(\beta)$$

og trekker man disse fra hverandre, får man

$$\cos(\alpha-\beta) - \cos(\alpha+\beta) = 2\sin(\alpha)\sin(\beta)$$

Dersom $m = n$, kan vi gjøre slik:

$$\int_{-T/2}^{T/2} \sin^2(n\omega t)\, dt = \frac{1}{2}\int_{-T/2}^{T/2} 1 - \cos(2n\omega t)\, dt = \frac{T}{2}$$

og dersom $n \neq m$, kan vi gjøre slik

$$\int_{-T/2}^{T/2} \sin(n\omega t)\sin(m\omega t)\, dt = \frac{1}{2}\int_{-T/2}^{T/2} \cos((n-m)\omega t) - \cos((n+m)\omega t)\, dt = 0$$

De andre beregningene er tilsvarende. Jeg kunne fortsatt med å gange med $\sin m\omega t$ og integrert og kansellert, men helt seriøst, det er akkurat det samme som i beviset for fun fact 4 over. Du må bare forstå hva et indreprodukt er og så er det gudd.

$\boxed{4}$ Du kan sjekke at

$$(f, g) = \sum_{k=1}^{n} f(x_k)\, g(x_k),$$

er et indreprodukt, og når vi finner minste kvadrats løsning, minimerer vi normen i akkurat dette indreproduktet. Fun fact fem sier at dette får du til nettopp ved å velge fourierkoeffisientene. Å måle feilen ved L^2-normen slik

$$\|x-y\|^2 = \int_{-T/2}^{T/2} (x(t) - y(t))^2\ dt$$

istedet, prøver du på en måte å minimere avstanden mellom x og y overalt istedet for bare i et endelig antall punkter, men prinsippet er det samme.

$\boxed{6}$ Dette er oppgave 6 herfra:
`https://www.math.ntnu.no/emner/TMA4110/2018h/eksamen/tma4110-2018h-nb.pdf` og du finner løsningen her:
`https://www.math.ntnu.no/emner/TMA4110/2018h/eksamen/tma4110-2018h-lf.pdf`

$\boxed{7}$ Hvis du leser om legendrepolynomene på wikipedia, vil du se at de skal være ortogonale med hensyn på indreproduktet

$$\int_{-1}^{1} f(x)g(x)\, dx.$$

De to første moniske polynomene $P_0(x) = 1$ og $P_1(x) = x$ er allerede ortogonale med hensyn på indreproduktet over. Det neste, x^2, er ortogonal på P_1, men ikke på P_0. Vi prøver Gram-Schmidt, og får

$$\begin{aligned} P_2(x) &= x^2 - \frac{\int_{-1}^{1} x^2 \cdot x \, dx}{\int_{-1}^{1} x \cdot x \, dx} \cdot x - \frac{\int_{-1}^{1} x^2 \cdot 1 \, dx}{\int_{-1}^{1} 1 \cdot 1 \, dx} \cdot 1 \\ &= x^2 - \frac{0}{2/3} x + \frac{2/3}{2} \\ &= x^2 - \frac{1}{3} \end{aligned}$$

Merk ellers hvor deilig det er at $\cdot$ er frigjort til vanlig multiplikasjon etter at vi begynte å skrive $\mathbf{x}^T\mathbf{y}$ istedet for $\mathbf{x} \cdot \mathbf{y}$. Det er vanlig å skalere legendrepolynomene slik at $P_n(1) = 1$, og $P_2(1) = 2/3$. Derfor ganger vi P_2 med $3/2$. Fortsetter vi slik som dette, får vi polynomene

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \frac{1}{2}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3) \\ &\vdots \end{aligned}$$

Det går an å vise at de tilfredsstiller rekursjonen

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

men det skal vi ikke gjøre.

$\boxed{8}$ Hvordan man løser denne avhenger litt av hvordan man tenker på uttrykket

$$(f, g) = \sum_{k=1}^{n} f(x_k)\, g(x_k)\,.$$

Hvis du tenker på f og g som diskrete signaler, kun definert i gitterpunktene x_k, blir det veldig greit, for da er uttrykket over bare det vanlige skalarproduktet, og åpenbart et indreprodukt siden skalarproduktet er et indreprodukt.

Dersom du tenker på f og g i indreproduktet som funksjoner på $\mathbb{R}$ som er samplet i gitterpunktene, må vi være litt forsiktige, for aksiom nummer to sier at

$$(\mathbf{x}, \mathbf{x}) > 0 \quad \text{dersom} \quad \mathbf{x} \neq \mathbf{0}.$$

Hvis du ser nøye på hvordan dette aksiomet er formulert, ser du at det inneholder en implikasjon om at dersom $(\mathbf{x}, \mathbf{x}) = 0$, må $\mathbf{x}$ være nullvektoren. At nullvektoren må ha lengde null følger av det tredje aksiomet, siden

$$(\mathbf{0}, \mathbf{0}) = (0 \cdot \mathbf{x}, 0 \cdot \mathbf{x}) = 0 \cdot 0 \cdot (\mathbf{x}, \mathbf{x}) = 0,$$

og det andre aksiomet sier altså at ingen andre vektorer kan ha lengde null. Men dersom du tar en funksjon f som er har et nullpunkt i alle gitterpunkter, vil

$$(f, f) = \sum_{k=1}^{n} f(x_k)^2 = 0$$

selv om f ikke er konstant lik null, så vi ser nå at "litt" av aksiom to ikke holder. Men det er ikke så veldig alvorlig, for mye som er sant for indreprodukter, er fortsatt sant om dette aksiomet slakkes litt på. Fun factene i denne økten er for eksempel sanne dersom man er nøye på å kreve at den ovennevnte funksjonen holdes utenfor; dette kan du få til ved å kreve at basiser skal være ortonormale og ikke bare ortogonale. Nå er det ikke første gang i verdenshistorien noen støter på noe som er av interesse men som bare nesten tilfredsstiller det andre aksiomet, så det har et navn, og det er **semi-indreprodukt**,

$\boxed{9}$ Alt dette betyr at vi kan minimere avstanden mellom responsvariabelen og forklaringsvariabelen med avstanden indusert av indreproduktet over, og forvente at fourierkoeffisientene gir oss den korrekte minimerte avstanden. La oss tenke at vi har et datasett

$$(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix}$$

og la $y(x)$ være en funksjon med funksjonsverdier $\mathbf{y}$. Vi er interessert i å finne den rette linjen $p(x) = \beta_1 x + \beta_0$ som minimerer avstanden til $y(x)$ målt i indreproduktet, som blir:

$$(p_1 - y, p_1 - y) = \sum_{k=1}^{n} (\beta_1 x_k + \beta_0 - y_k)^2$$

Dette er akkurat den kvadratsummen vi minimerte i forrige uke. En basis for alle rette linjer er $\{1, x\}$, men disse er ikke ortogonale med hensyn på dette indreproduktet, så vi må ortogonalisere litt. La oss kalle den ortogonale basisen $\{v_0, v_1\}$ og sette $v_1 = 1$. Gram-Schmidt gir

$$v_1(x) = x - \frac{\sum_{k=1}^{n} x_k \cdot 1}{\sum_{k=1}^{n} 1 \cdot 1} = x - \overline{\mathbf{x}},$$

og basisen gitt av $v_0(x) = 1$ og $v_1(x) = x - \overline{\mathbf{x}}$ er ortogonal, så nå er det bare å bruke fun fact 5, og skrive

$$p_1 = \frac{(y, v_0)}{(v_0, v_0)} v_0 + \frac{(y, v_1)}{(v_1, v_1)} v_1.$$

Siden

$$\frac{(y, v_0)}{(v_0, v_0)} v_0 = \frac{\sum_{k=1}^{n} y_k \cdot 1}{\sum_{k=1}^{n} 1 \cdot 1} = \overline{\mathbf{y}}$$

og

$$\frac{(y, v_1)}{(v_1, v_1)} v_1 = \frac{\sum_{k=1}^{n} y_k \cdot (x_k - \overline{\mathbf{x}})}{\sum_{k=1}^{n} (x_k - \overline{\mathbf{x}}) \cdot (x_k - \overline{\mathbf{x}})} = \overline{\mathbf{y}} x$$

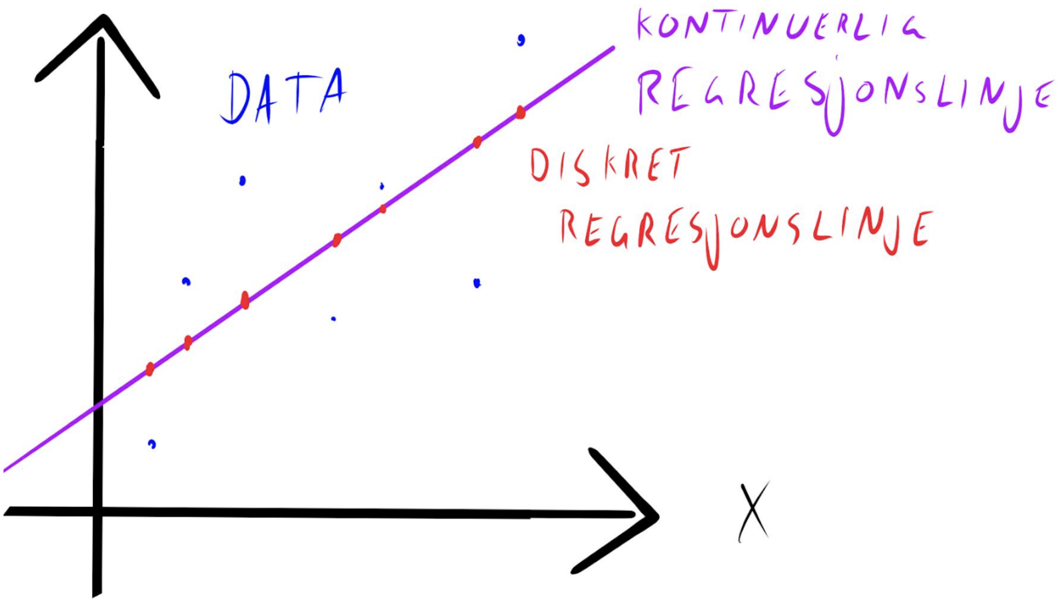
$$= \frac{\mathbf{x}^T \mathbf{y} - n \overline{\mathbf{x}\mathbf{y}}}{\|\mathbf{x}\|^2 - n(\overline{\mathbf{x}})^2}$$

får vi

$$p_1(x) = \overline{\mathbf{y}} + \frac{\mathbf{x}^T\mathbf{y} - n\overline{\mathbf{x}\mathbf{y}}}{\|\mathbf{x}\|^2 - n(\overline{\mathbf{x}})^2}\,(x - \overline{\mathbf{x}})$$

som du kan dobbeltsjekke er rett svar.

Det som er forskjellen på de to forskjellige tolkningene av indreproduktet, er hvorvidt denne løsningen er entydig. Hvis du tenker kontinuerlig, vil det ikke finnes noen entydig funksjon om minimerer avstanden, for du kan legge til en vilkårlig funksjon med nullpunkter i alle x_k uten at avstandsmålet "ser" noe til denne. Dersom du tenker diskret, får du ikke dette problemet. Men jeg har ikke plaget dere med entydighet, så ikke tenk så mye på det.



$\boxed{10}$ La oss begynne med å lage en (for vårt bruk) bedre ortonormal basis for $\mathbb{R}^3$. La

$$\mathbf{v}_1 = \mathbf{y} \qquad \mathbf{v}_2 = \frac{\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}}{\|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\|} \qquad \mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2 = \frac{\mathbf{y} \times \mathbf{x}}{\|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\|}$$

Trikset er nå å skrive $\mathbf{x}$ i denne basisen:

$$\mathbf{x} = (\mathbf{x}^T\mathbf{y})\,\mathbf{v}_1 + \big\|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\big\|\,\mathbf{v}_2 + 0 \cdot \mathbf{v}_3$$

Dette ser mer komplisert ut, men nå kan vi rotere de to siste koordinatene

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\| \\ 0 \end{pmatrix} = \|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\| \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

og få

$$\begin{aligned} \text{rotert } \mathbf{x} &= (\mathbf{x}^T\mathbf{y})\,\mathbf{v}_1 + \big\|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\big\|\,\mathbf{v}_2 \cos\theta + \big\|\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y}\big\|\,\mathbf{v}_3 \sin\theta \\ &= (\mathbf{x}^T\mathbf{y})\,\mathbf{y} + (\mathbf{x} - (\mathbf{x}^T\mathbf{y})\mathbf{y})\cos\theta + \mathbf{y} \times \mathbf{x} \sin\theta \\ &= \mathbf{y}\,(\mathbf{x}^T\mathbf{y})\,(1 - \cos\theta) + \mathbf{x}\cos\theta + \mathbf{y} \times \mathbf{x} \sin\theta. \end{aligned}$$

$\boxed{12}$ Du får identitetsmatrisen! For ortonormale matriser er $B^T = B^{-1}$. Dette skal vi bruke hardt kommende uke.

$\boxed{13}$ Vi kan faktisk klare enda bedre. La

$$\mathbf{x} = \sum_{k=1}^{n} c_k \mathbf{v}_k \qquad \text{er} \qquad \mathbf{y} = \sum_{k=1}^{n} d_k \mathbf{v}_k$$

Da er

$$\begin{aligned}(\mathbf{x}, \mathbf{y}) &= \left(\sum_{k=1}^{n} c_k \mathbf{v}_k, \sum_{l=1}^{n} d_l \mathbf{v}_l \right) \\ &= \sum_{k=1}^{n} \sum_{k=1}^{n} c_k d_k \left(\mathbf{v}_k, \mathbf{v}_l \right) = \sum_{k=1}^{n} c_k d_k\end{aligned}$$

Likningen

$$(\mathbf{x}, \mathbf{x}) = \sum_{k=1}^{n} c_k^2$$

er et spesialtilfelle. I $\mathbb{R}^2$ ser det slik ut: