MAT1300 – Analyse I – 23. februar 2009

La $E \subset R^m$ og $f \: E \to R^p$. Anta $B(x, \delta) \subset E$. Dersom f er deriverbar i x, er den deriverte $\alpha = f'(x) \: R^m \to R^p$ gitt ved multiplikasjon med en p x m matrise $A = (a_{i,j})_{i,j=1,1}^{p,m}$. Vi kan tenke på f som et p-tuppel av funksjoner

$$f_1, \dots, f_p \: E \to R$$

der hver $f_i \: E \to R$ er deriverbar i x, for $1 \le i \le p$. Vi kan skrive likningen

$$(*) \qquad f(x+h) = f(x) + \alpha(h) + \epsilon(x,h)||h||$$

i R^p som et p-tuppel av likninger

$$(**) \qquad f_i(x+h) = f_i(x) + \alpha_i(h) + \epsilon_i(x, h)||h||$$

i R, der

$$f(x) = (f_1(x), \dots, f_p(x))$$

$$\alpha(h) = (\alpha_1(h), \dots, \alpha_p(h))$$

$$\epsilon(x, h) = (\epsilon_1(x, h), \dots, \epsilon_p(x, h)) .$$

Spesielt er $\alpha_i \: R^m \to R$ en lineæravbildning for hver i, gitt ved multiplikasjon med 1 x m matrisen $(a_{i,j})_{j=1}^m$ gitt ved i-te rad i A. Vi kan derfor bestemme matrisen A og lineæravbildningen $\alpha \: R^m \to R^p$, ved å bestemme hver rad i A og hver lineæravbildning $\alpha_i \: R^m \to R$.
Siden $\epsilon(x, h) \to 0$ hvis og bare hvis hver $\epsilon_i(x, h) \to 0$, for $1 \le i \le p$, er f deriverbar i x hvis og bare hvis hver f_i er deriverbar i x, og ved likningen (**) er $\alpha_i = f'_i(x)$ i så fall lik den deriverte til f_i i x.

Vi fokuserer derfor på en enkelt funksjon $g \: E \to R$, som er deriverbar i x. Den deriverte $\beta = g'(x) \: R^m \to R$ av g i x er bestemt av likningen

$$g(x+h) = g(x) + \beta(h) + \epsilon(x, h)||h||$$

der $\epsilon(x, h) \to 0$ når $h \to 0$. Her er $h \in R^m$, med $||h||<\delta$. For å redusere til en reell funksjon av {\'e}n reell variabel, der vi kan beskrive den deriverte som grensen av en differential-kvotient, begrenser vi h til å ligge på en linje gjennom origo. La $u \in R^m$ være en enhetsvektor, slik at $||u||=1$. For eksempel kan $u = e_j$, der $1 \le j \le m$. Funksjonen

$$G(t) = g(x + tu)$$

er da en funksjon $G \: (-\delta, \delta) \to R$, som implisitt avhenger av u, med

$$G(t) = G(0) + \beta(tu) + \epsilon(x, tu) |t| .$$

Her er $t \mapsto \beta(tu)$ en lineæravbildning $R \to R$, så det finnes et entydig bestemt tall $b \in R$ med $\beta(tu) = bt$ for alle t. Altså er

$${G(t) - G(0) \over t} = b +/- \epsilon(x, tu) .$$

Dersom g er deriverbar i x vet vi at $\epsilon(x, tu) \to 0$ når $t \to 0$, så G er deriverbar i $t=0$, med derivert $G'(0) = b = \beta(u)$.

Omvendt, dersom G er deriverbar i $t=0$, så sier vi at g har en retningsderivert i x i retningen u, med verdi $G'(0)$.

For $u = e_j$, der $1 \le j \le m$, kaller vi den retningsderiverte til g i x i retningen e_j den j-te partielle deriverte av g i x, dersom den eksisterer. Dette er grensen

$$G'(0) = \lim_{t\to0} {g(x+t e_j) - g(x) \over t} .$$

Her er $(x + te_j) = (x_1, \dots, x_{j-1}, x_j + t, x_{j+1}, \dots, x_m)$. Vi skriver

$$D_j g(x) = g_{,j}(x) \text{ eller } {\partial g \over \partial x_j}(x)$$

for denne retningsderiverte.

Dersom g er deriverbar i x har den altså retningsderiverte i x i alle retninger u, og spesielt har den j-te partielle deriverte i x for alle $1 \le j \le m$. Det omvendte gjelder generelt ikke: det finnes funksjoner med retningsderiverte i alle retninger som ikke er deriverbare. Se Example 7.32.

Lineæravbildningen $\beta \: R^m \to R$ er bestemt av sine verdier

$$b_j = \beta(e_j)$$

på basisvektorene e_j. Da er

$$\beta(h) = b_1 h_1 + \dots + b_m h_m = \sum_{j=1}^m b_j h_j$$

for $h = (h_1, \dots, h_m)$, så $(b_j)_{j=1}^m$ er matriseformen til β.

Når g er deriverbar i x med $g'(x) = \beta$ har vi sett at den j-te partielle deriverte $D_j g(x) = g_{,j}(x)$ er lik $G'(0) = \beta(e_j)$, der $G(t) = g(x + t e_j)$. Altså er $D_j g(x) = g_{,j}(x) = b_j$, så matriseformen til $Dg(x) = g'(x) = \beta$ er gitt ved de partielle deriverte

$$(D_j g(x))_{j=1}^m = (g_{,j}(x))_{j=1}^m .$$

Vi vender så tilbake til $f \: E \to R^p$, med koordinater $f_i \: E \to R$ for $1 \le i \le p$. Dersom f er deriverbar i x er hver f_i deriverbar i x, med partielle deriverte

$$D_j f_i(x) = f_{i,j}(x) = {\partial f_i \over \partial x_j}(x)$$

for hver j. Matriseformen til $\alpha = Df(x) = f'(x)$ har p rader, med i-te rad lik matriseformen til $Df_i(x) = f'_i(x)$, og denne er lik $(D_j f_i(x))_{j=1}^m = (f_{i,j}(x))_{j=1}^m$. Altså er matriseformen til $Df(x) = f'(x)$ lik $p \times m$ matrisen

$$(D_j f_i(x))_{i,j=1,1}^{p,m} = (f_{i,j}(x))_{i,j=1,1}^{p,m} .$$

Dette kalles Jacobi-matrisen til f i x, og består av alle de partielle deriverte av alle koordinatene til f.

Lemma: La $E \subset R^m$, $f \colon E \to R^p$, $B(x, \delta) \subset E$ og anta at f er deriverbar i x. Da eksisterer alle de partielle deriverte $f_{i,j}(x)$, og matrisen til den deriverte $Df(x) \colon R^m \to R^p$ er Jacobi-matrisen $(f_{i,j}(x))_{i,j=1,1}^{p,m}$.

6.2. Operatornormen

Vi trenger et mål på størrelsen av en lineæravbildning $\alpha \colon R^m \to R^p$.

Lemma:
Hvis $\alpha \colon R^m \to R^p$ er lineær finnes en konstant $K = K(\alpha)$ slik at

$$||\alpha(x)|| \le K ||x||$$

for alle $x \in R^m$.

Bevis:
La $A = (a_{i,j})$ være matriseformen til α, slik at for $x = (x_1, \dots, x_m)$ er $y = \alpha(x) = (y_1, \dots, y_p)$ gitt ved

$$y_i = \sum_{j=1}^m a_{i,j} x_j$$

for hver $1 \le i \le p$. La $K = \sum_{i,j} |a_{i,j}|$. Da er

$$||y|| \le \sum_i |y_i| \le \sum_{i,j} |a_{i,j}| |x_j|$$

$$\le \sum_{i,j} |a_{i,j}| ||x|| \le K ||x|| .$$

Alternativt bevis:
La $B = B(0, 1) = \{ x \in R^m : ||x|| \le 1\}$. Da er B lukket, begrenset og ikketom. Funksjonen $f \colon B \to R$ gitt ved

$$f(x) = ||\alpha(x)||$$

er kontinuerlig, og oppnår derfor sitt maksimum. Med andre ord finnes en m in B med $f(y) \le f(m)$ for alle $y \in B$. La $K = f(m)$. For en generell $x \in R^m - \{0\}$ er $y = x/||x|| \in B$, så

$$K \ge f(y) = ||\alpha(x/||x||)|| = ||\alpha(x)||/||x||$$

som impliserer $||\alpha(x)|| \le K ||x||$. For $x=0$ er dette klart. QED.

Definisjon:
La $\alpha \: R^m \to R^p$ være lineær. La operatornormen

$$||\alpha|| = \sup_{||x|| \le 1} ||\alpha(x)||$$

være supremum av mengden av normer $||\alpha(x)||$ der $||x|| \le 1$.

Ved det første beviset ovenfor er denne mengden oppad begrenset, og har derfor et supremum. Ved det andre beviset har mengden til og med et (oppnådd) maksimum.

Oppgave:

$$||\alpha|| = \sup_{||x|| \le 1} ||\alpha(x)||$$

$$= sup_{||x|| = 1} ||\alpha(x)||$$

$$= sup_{x \ne 0} { ||\alpha(x)|| \over ||x|| } .$$

Lemma: La $\alpha, \beta \: R^m \to R^p$ være lineæravbildninger.

(i) Hvis $x \in R^m$ er $||\alpha(x)|| \le ||\alpha|| ||x||$.

(ii) $||\alpha|| \ge 0$.

(iii) Hvis $||\alpha|| = 0$ så er $\alpha=0$.

(iv) Hvis $k \in R$ er $||k \alpha|| = |k| ||\alpha||$.

(v) (trekantulikheten)

$$||\alpha + \beta|| \le ||\alpha|| + ||\beta|| .$$

(vi) Hvis $\gamma \: R^p \to R^q$ er lineær, så er

$$||\gamma \alpha|| \le ||\gamma|| ||\alpha|| .$$

Bevis av (v)
:For alle $x \in R^m$ er

$$||(\alpha + \beta)(x)|| = ||\alpha(x) + \beta(x)||$$

$$\le ||\alpha(x)|| + ||\beta(x)|| \le ||\alpha|| ||x|| + ||\beta|| ||x||$$

$$= (||\alpha|| + ||\beta||) ||x|| .$$

Dersom $||x|| \le 1$ er dette oppad begrenset av $||\alpha|| + ||\beta||$, så supremum av alle disse uttrykkene, dvs. $||\alpha + \beta||$, er mindre enn eller lik $||\alpha|| + ||\beta||$. QED.

La nå $f \: R^m \to R^p$ og $g \: R^p \to R^q$ være deriverbare. La $x \in R^m$ og sett $y = f(x)$. For små $h \in R^m$ er da $f(x+h) \approx f(x) + f'(x)(h)$, mens for små $k \in R^p$ er $g(y+k) \approx g(y) + g'(y)(k)$. Da er antageligvis

$$(gf)(x+h) = g(f(x+h)) \approx g(f(x) + f'(x)(h)) = g(y+k)$$

med $k = f'(x)(h)$, og

$$g(y+k) \approx g(y) + g'(y)(k) = gf(x) + g'(f(x))(f'(x)(h))$$

så det er rimelig å forvente at

$$(gf)'(x) = g'(f(x)) f'(x)$$

er den sammensatte lineæravbildningen av $f'(x) \: R^m \to R^p$ og $g'(f(x)) \: R^p \to R^q$.

Lemma (Kjerneregelen): La $U \subset R^m$, $V \subset R^p$, $f \: U \to R^p$ og $g \: V \to R^q$ slik at $f(U) \subset V$. La $x \in U$ og $y = f(x) \in V$. Anta at U er en omegn om x, og at V er en omegn om y. Anta at f er deriverbar i x med derivert $\alpha = f'(x) \: R^m \to R^p$ og at g er deriverbar i y med derivert $\beta = g'(y) \: R^p \to R^q$. Da er den sammensatte funksjonen $gf \: U \to \R^q$ deriverbar i x med derivert $\beta\alpha \: R^m \to R^q$. Med andre ord:

$$D(gf)(x) =Dg(f(x)) Df(x)$$

$$(gf)'(x) = g'(f(x)) f'(x) .$$

Bevis:
Vi vet at

$$f(x+h) = f(x) + \alpha(h) + \epsilon_1(h) ||h||$$

og

$$g(f(x)+k) = g(f(x)) + \beta(k) + \epsilon_2(k) ||k||$$

der $\epsilon_1(h) \to 0$ når $h \to 0$, og $\epsilon_2(k) \to 0$ når $k \to 0$. Det følger at

$$(gf)(x+h) = g(f(x+h)) = g(f(x) + \alpha(h) + \epsilon_1(h) ||h||)$$

så vi lar $k = \alpha(h) + \epsilon_1(h) ||h||$ og får

$$(gf)(x+h) = g(f(x)) + \beta(\alpha(h) + \epsilon_1(h) ||h||)$$

$$+ \epsilon_2(\alpha(h) + \epsilon_1(h) ||h||) ||\alpha(h) + \epsilon_1(h) ||h|| ||$$

$$= (gf)(x) + (\beta\alpha)(h) + \eta(h) ||h||$$

der

$$\eta(h) = \eta_1(h) + \eta_2(h)$$

og

$$\eta_1(h) = \beta(\epsilon_1(h))$$

mens

$$\eta_2(h) ||h||$$

$$= \epsilon_2(\alpha(h) + \epsilon_1(h) ||h||) ||\alpha(h) + \epsilon_1(h) ||h|| || .$$

Vi må vise at $\eta(h) \to 0$ når $h \to 0$. Det er nok å vise at $\eta_1(h) \to 0$ og $\eta_2(h) \to 0$ når $h \to 0$.

Vi vet at

$$||\eta_1(h)|| = ||\beta(\epsilon_1(h))|| \le ||\beta|| ||\epsilon_1(h)||$$

og $\epsilon_1(h) \to 0$ når $h \to 0$, så $\eta_1(h) \to 0$ når $h \to 0$.

Videre er

$$|| \alpha(h) + \epsilon_1(h) ||h|| || \le ||\alpha(h)|| + ||\epsilon_1(h)|| ||h||$$

$$\le ||\alpha|| ||h|| + ||\epsilon_1(h)|| ||h||$$

$$= (||\alpha|| + ||\epsilon_1(h)||) ||h||$$

så

$$||\eta_2(h)||$$

$$\le ||\epsilon_2(\alpha(h) + \epsilon_1(h) ||h||)|| (||\alpha|| + ||\epsilon_1(h)||)$$

$$\to 0 (||\alpha|| + 0) = 0$$

når $h \to 0$. QED.