1 Sammandrag av föreläsning 1 Torsdag 2/9/99

Inledning till rummet $\mathbf{R}^n$:

Vi betecknar som $\mathbf{R}^n$ mangden av n-tupler av reella tal, i.e.:

$$\mathbf{R}^n = \{(x_1, ..., x_n) | x_i \in \mathbf{R},\ i = 1, ..., n\}. \tag{1}$$

Denna mangd har tva egenskap som vi ar interesserad av :

1. Den ar ett *algebraiskt vektor rum*, med addition och skalarmultiplikation definierade som foljer :

$$(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n) \tag{2}$$

$$\alpha(x_1, ..., x_n) = (\alpha x_1, ..., \alpha x_n) \quad (\alpha \in \mathbf{R}. \tag{3}$$

2. Den ar ett *topologiskt metriskt rum*, i.e.: for varje tva element $x, y \in \mathbf{R}^n$, det finns ett icke-negativt reellt tal, betecknad $d(x, y)$ och definierad som

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}. \tag{4}$$

$d(x, y)$ kallas for *avstandet* mellan x och y.

Forhallandet mellan egenskap **1** och **2** ar att funktionen d kan definieras i termer av en *skalarprodukt* pa vektorrummet. Denna ar en funktion $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$, definierad som

$$x \cdot y = \sum_{i=1}^{n} x_i y_i. \tag{5}$$

Vi kan definiera *langden* av en vektor som

$$|x| = \sqrt{x \cdot x} \tag{6}$$

och da ser vi att

$$d(x, y) = |x - y| \quad \text{for alla } x, y \in \mathbf{R}^n. \tag{7}$$

Obs : Ett vektor rum med en skalarprodukt kallas for en *inner product space*. I ett sadant rum, man far alltid en mening for 'avstandet mellan tva punkter', genom formler (6) och (7).

Definition 1 : En *reellvard funktion av n reella variabler* ar en funktion $f : D \to R$, dar D ar en delmangd av $\mathbf{R}^n$.

I fjol, ni studerade fallet $n = 1$. I matematisk analys, och i de flesta tillampningarna, man ar interesserad av *kontinuerliga* och *differentierbara* funktioner. I fallet $n = 1$ dessa begrepp var ganska klart att forsta (hoppas jag !!). I fallet $n > 1$, kontinuitet ar fortfarande ganska latt att forklara (det gor vi nasta gang), men differentierbarhet ar ett mycket mer subtilt begrepp. Mer senare .....

Definition 2 : En *vektorvard reell funktion av n reella variabler* ar en funktion $f : D \to \mathbf{R}^m$, dar $D \subseteq \mathbf{R}^n$ och $m \geq 1$.

(I fallet $m = 1$, far vi tillbaka definition 1).

I manga tillampningar, de funktionerna som upptrader ar vektorvarda. Men, fran den teoretiska synpunkten, sadana funktioner ar inte mer komplicerada an reellvarda funktioner, ty en v.v. funktion $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^m$ bestamms av (?? svenska : eng. = is determined by) m r.v. funktioner, namligen dess *komponent funktioner* f_i, $i = 1, ..., m$, dar

$$f(x) = (f_1(x), ..., f_m(x)) \qquad (x \in D). \tag{8}$$

Geometrisk uppfattning :

Nar n och m ar bada ≤ 3, vi far en geometrisk uppfattning till v.v. funktioner.

$n = 1$:

En sadan funktion kallas for en *kurva* (i m-space). En kurva i 1-space ar en vanlig reellvard funktion av en variabel. En kurva i 2-space kallas for en *plan kurva/ kurva i planet* (eng.: plane curve). Markera att grafen av en vanlig en-variabel funktion f ar en plan kurva $F : \mathbf{R} \rightarrow \mathbf{R}^2$, dar

$$F(t) = (t, f(t)). \tag{9}$$

Daremot (eng.: conversely), en plan kurva ar grafen av en vanlig en-variabel funktion omm den skarnar varje vertikal linje i hogst en punkt.

En kurva i 3-space kallas for en *rymd kurva/ kurva i rymden* (eng. = space curve). Nu visar vi nagra exempel :

Exempel 1 : Betrakta rymd kurvan f dar

$$f(t) = (x_0 + \alpha t, y_0 + \beta t, z_0 + \gamma t). \tag{10}$$

Den ar den ratta linjen genom (x_0, y_0, z_0) och parallell med vektorn $\alpha \mathbf{e}_x + \beta \mathbf{e}_y + \gamma \mathbf{e}_z$.

Exempel 2 : Rymd kurvan $f : \mathbf{R}^+ \rightarrow \mathbf{R}^3$ dar

$$f(t) = (t \cos t, t \sin t, t) \tag{11}$$

ar en spiral kring z-axeln.

$n = 2$:

En sadan funktion kallas for en *yta*. Vi ar interesserad av ytor i rymden. Markera att grafen av en reellvard funktion f av 2 variabler ar rymdytan F dar

$$F(s, t) = (s, t, f(s, t)). \tag{12}$$

En sadan yta skarnar varje linje parallell med z-axeln i hogst en punkt. Men de flesta interessanta ytor har inte detta egenskap. Vi tar, till exempel, en sfar.

Exempel : Lat S beteckna sfaren med centrum $(0, 0, 0)$ och radie r. Vi skapar en *bijektion* $f : [0, \pi) \times [0, 2\pi) \rightarrow S$ genom att satta

$$f(\theta, \phi) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta). \tag{13}$$

For att se att f ar en bijektion pa S, titta pa bilden pa sidan 22 i boken. Denna ar den vanligaste beskrivningen av en sfar som en rymdyta.

OBS : Fran forgaende exempel, ser vi att for varje punkt $(x, y, z) \in \mathbf{R}^3$, det existerar en unik tripel (r, θ, ϕ) av reella tal, dar $r \geq 0$, $0 \leq \theta < \pi$, $0 \leq \phi < 2\pi$, sa att

$$x = r \sin\theta \cos\phi \tag{14}$$

$$y = r \sin\theta \sin\phi \tag{15}$$

$$z = r \cos\theta. \tag{16}$$

r, θ och ϕ kallas for *sfariska koorinater* i rymden. Nasta gang, ska vi prata mer om koordinat system.

Andra interessanta kurvor och ytor diskuteras i boken, t.ex. de *kvadratiska* kurvorna och ytorna.

2 Sammandrag av forelasning 2
Tisdag 7/9/99

De viktiga begrepp i, och tillampningar av, flervariabelanalys ar, i stortsatt, desamma som i det 1-variabel fallet.

Viktiga begrepp :

1. Gransvarde och kontinuitet.
2. Deriverbarhet.
3. Integration av kontinuerliga funktioner.

Viktiga tillampningar :

4. Optimiserings problem.
5. Losning av diffekvationer.

I denna kurs (del 1 plus del 2), kommer vi att generalisera dessa begrepp och tillampningar till hogre dimensioner.

Gransvarde och kontinuitet

Kom ihag foljande definitioner fran 1-variabel analys :

DEFINITION 1 : Lat $f : D \subseteq \mathbf{R} \rightarrow \mathbf{R}$ och $a \in \mathbf{R}$. Vi sager att f har *gransvardet* $L \in \mathbf{R}$ som x gar mot a, och skriver

$$\lim_{x \rightarrow a} f(x) = L \tag{17}$$

om *(1)* f ar definierad i ett oppet intervall kring $x = a$ (kanske med undantag av $x = a$) och *(2)* for varje $\epsilon > 0$, det existerar $\delta > 0$ sa att

$$0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon. \tag{18}$$

DEFINITION 2 : Lat $f : D \subseteq \mathbf{R} \rightarrow \mathbf{R}$, och $a \in D$. Vi sager att f ar *kontinuerlig* vid $x = a$ om

$$\lim_{x \rightarrow a} f(x) = f(a). \tag{19}$$

OBS : Def. 1 $\Rightarrow$ om f ar kontinuerlig vid $x = a$, da maste f vara definierad i ett oppet intervall kring $x = a$.

For att utviga de har definitionerna till hogre dimensioner, maste vi generalisera tva begrepp, namligen :

(a) belopp av ett reellt tal
(b) oppet intervall i $\mathbf{R}$.

Vi har redan tagit hand om (a) forra gangen, dar vi definierade beloppet av en vektor i $\mathbf{R}^n$. Med avseende pa (b), den motsvarande grejen i $\mathbf{R}^n$ kallas for ett *oppet klot*, som vi definierar nu :

DEFINITION 3 : Lat $a \in \mathbf{R}^n$ och $r > 0$ vara ett positivt reellt tal. Det *oppna klotet* med centrum a och radie r, betecknad $B(a,r)$, ar mangden av alla punkter i $\mathbf{R}^n$ vilkas avstand fran a ar mindre an r, i.e.:

$$B(a,r) = \{x \in \mathbf{R}^n : |x - a| < r\}. \tag{20}$$

OBS : En delmangd av $\mathbf{R}^n$ som innehaller ett oppet klot kring var och en av dess punkter kallas for en *oppen mangd*, i.e.: $D \subseteq \mathbf{R}^n$ ar en oppen mangd om, for varje $a \in D$, det existerar $r > 0$ (beroende av a) sa att $B(a,r) \subseteq D$.

Nu kan vi generalisera def. 1 och 2 till (vektorvarde) funktioner av flera variabler :

DEFINITION 4 : Lat $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^m$ och $a \in \mathbf{R}^n$. Vi sager att f har *gransvardet* $L \in \mathbf{R}^m$ som x gar mot a i $\mathbf{R}^n$, och skriver

$$\lim_{x \to a} f(x) = L \tag{21}$$

om *(1)* f ar definierad i ett oppet klot kring $x = a$ (kanske med undantag av $x = a$) och *(2)* for varje $\epsilon > 0$, det existerar $\delta > 0$ sa att

$$0 < |x - a|_{\mathbf{R}^n} < \delta \Rightarrow |f(x) - L|_{\mathbf{R}^m} < \epsilon. \tag{22}$$

DEFINITION 5 : Lat $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^m$ och $a \in D$. Vi sager att f ar *kontinuerlig* vid $x = a$ om

$$\lim_{x \to a} f(x) = f(a). \tag{23}$$

OBS : Fran ovanstaende definitioner ser vi att, om f ar kontinuerlig vid alla punkter i dess definitionsmangd D, da maste D vara en oppen mangd i $\mathbf{R}^n$.

Var forsta proposition sager, vasentligen (essentially), att kontinuitet for vektorvarda funktioner kan reduceras till reellvarda funktioner. Mer precis

Proposition 1. *Med desamma beteckningar som i def. 4 och 5, antag att* $L = (l_1, ..., l_m)^T$ *och* $f(x) = (f_1(x), ..., f_m(x))^T$. *Da*

(1) $\lim_{x \to a} f(x) = L$ *omm* $\lim_{x \to a} f_i(x) = l_i$ *for varje* i, $1 \leq i \leq m$

(2) f *ar kontinuerlig vid* $x = a$ *omm varje* f_i *ar kontinuerlig vid* $x = a$.

BERAKNINGAR :

Genom resten av detta avsnitt, f betecknar en reellvard funktion av n variabler (for nagon $n \geq 1$).

Om begreppen gransvarde/kontinuitet ar ganska latt att forsta, de ar inte alltid sa latt att jobba med. Den allmana fragan som vi vill bearbeta ar

For en given funktion $f : \mathbf{R}^n \to \mathbf{R}$ *och en given punkt* $a \in \mathbf{R}^n$,

(a) hur kan man bestamma sig om $\lim_{x \to a} f(x)$ *existerar ?*
(b) om gransvardet existerar, hur beraknar man det ?

Det ar klart att del (a) av detta problem ar den viktigaste, ty den handlar om att BEVISA nagonting. Om vi kan bevisa att gransvardet inte existerar, da ar problemet lost. Om vi kan bevisa motsatsen, beviset brukar ge gransvardet (t.ex. 1.27) och, aven om det inte ar sa, vi kan approximera gransvardet med hjalp av en dator.

FRAGA 1 : Om man tror att $\lim_{x \to a}$ inte existerar, hur kan man bevisa den ?

SVAR : Det finns ingen allman metod, men den vanligaste ideen ar att 'reducera' problemet till det 1-dim. fallet. Punkten ar att om f ar inskrankt (restricted) till en kurva, den blir en funktion av en variabel. Da ar det klart att $\lim_{x \to a} f(x)$ inte existerar om antingen ... eller

(1) det finns en kurva genom $x = a$ sa att, langs kurvan, f har inget gransvarde vid $x = a$

Exempel 1 : $\lim_{(x,y)\to(0,0)} \sin(\frac{y}{x^2})$ existerar inte ty, langs kurvan $y = x$ funktionen blir $\sin(1/x)$, som har inget gransvarde vid $x = 0$.

(2) det finns tva olika kurvor genom a sa att, langs var och en kurva f har ett gransvarde, men de tva gransvarden ar olika.

Exempel 2 : $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ existerar inte ty, langs den ratta linjen $y = \alpha x$ genom $(0,0)$, gransvardet ar $\frac{\alpha}{1+\alpha^2}$, som ger olika varde for olika α.

OBS : I (2), de enklaste kurvor att jobba med ar de ratta linjerna genom $x = a$. Men aven om gransvarden existerar och ar lika med varandra langs alla sadana linjer, det kan fortfarande handa att f har inget gransvarde :

Exempel 3 : $\lim_{(x,y)\to(0,0)} \frac{x^4y^4}{x^6+y^{12}}$ existerar inte. Langs varje ratt linje $y = \alpha x$ har vi gransvardet 0, men langs kurvan $y = \sqrt{x}$ gransvardet $= \frac{1}{2}$.

FRAGA 2 : Om man tror att $\lim_{x\to a} f(x)$ existerar, hur kan man bevisa/berakna den ?

SVAR : En gang till, ingen allman metod. Det foljer fran Def. 4 att, OM f har ett gransvarde langs varje kurva genom a och alla sadana varde ar lika med varandra, sag $= L$, da ar $\lim_{x\to a} f(x) = L$. Men man kan inte kolla alla kurvor !! Du maste ha nagon 'intuition' i borjan. Da kan du (1) berakna gransvardet L langs, t.ex., en ratt linje och (2) forsoka (pa nagon satt) bevisa att $|f(x) - L| \to 0$ dar $|x - a| \to 0$.

Exempel 4 : (1.27 i ovningsboken)

$$\lim_{(x,y)\to(0,0)} \frac{x^4y}{x^2+(x+y)^2} = 0. \tag{24}$$

Beviset gavs pa forelasningen.

Dock, de flesta vanliga funktioner ar kontinuerliga. Dessa funktioner byggs upp fran (are built up from) elementara funktioner som ar solklart kontinuerliga. Som i det 1-dim fallet, det ar latt att bevisa att en summa, differens, produkt eller sammansattning av kontinuerliga funktioner ar kontinuerliga. Vi har foljande formuler :

$$\lim_{x \to a}(f(x) \pm g(x)) = \lim_{x \to a} f(x) \pm \lim_{x \to a} g(x) \tag{25}$$

$$\lim_{x \to a} f(x)g(x) = \lim_{x \to a} f(x) \times \lim_{x \to a} g(x) \tag{26}$$

$$\lim_{x \to a} f(g(x)) = \lim_{x \to g(a)} f(x). \tag{27}$$

(Samma bevis som i det 1-d fallet). Vi har ocksa att

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim_{x \to a} f(x)}{\lim_{x \to a} g(x)}, \tag{28}$$

forutsatt att $\lim_{x \to a} g(x) \neq 0$. Fran exempel 2,3,4 ser vi att svarigheter brukar uppstar nar man har ett sadant kvot men namnaren gar mot noll, exakt som i envariabelanalys.

Deriverbarhet

For att generalisera detta begrepp till hogre dimensioner, det ar bast att utnytta var geometrisk uppfattning i det 1-d fallet.

Obs. 1 : Geometriskt, en funktion $f : \mathbf{R} \to \mathbf{R}$ ar deriverbar vid $x = a$ om det finns en unik ratt linje tangent till grafen av f vid punkten a.

Obs. 2 : Ekvivalent (equivalently), f ar deriverbar vid a omm f kan approximeras med en ratt linje i narheten av a.

Nu kom ihag den precisa definitionen :

Definition : $f : \mathbf{R} \to \mathbf{R}$ ar deriverbar vid $x = a$ om

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{29}$$

existerar.

Antag nu att f ar deriverbar vid $x = a$ och lat gransvardet i (13) kallas for A. Da harledar (deduce) vi att, i narheten av a,

$$f(x) = f(a) + A(x-a) + \tau(x), \tag{30}$$

dar $\left|\frac{\tau(x)}{x-a}\right| \to 0$ som $x \to a$.

Det ar ekvation (30) som visar hur man skall ge en precis formulering av deriverbarhet i hogre dimensioner (def. B nere). Forst, en intuitiv formulering :

Definition A : Lat D vara en oppen mangd i $\mathbf{R}^n$, $f : D \to \mathbf{R}^m$ och $a \in D$. Vi sager att f ar deriverbar vid a om differensen $f(x) - f(a)$ kan approximeras med en linjar funktion i narheten av a.

Kom ihag fran linjar algebra att en linjar funktion mellan tva vektor rum kan beskrivas med en matris - det finns foljande sats :

Sats. *Lat $f : \mathbf{R}^n \to \mathbf{R}^m$ vara en linjar funktion, dvs en funktion sa att*

(1) $f(v_1 + v_2) = f(v_1) + f(v_2)$ for alla $v_1, v_2 \in \mathbf{R}^n$

(2) $f(\alpha v) = \alpha f(v)$ for alla $\alpha \in \mathbf{R}$, $v \in \mathbf{R}^n$.

Da, existerar det en $m \times n$ matris A sa att $f(x) = Ax$ for alla $x = (x_1, ..., x_n)^T \in \mathbf{R}^n$.

Nu har vi allt vi behover for att ge en precis formulering av deriverbarhet - men du maste vanta till nasta gang for att se den !!!

3 MAN 030 : Sammandrag av forelasning 3 Torsdag 9/9/99

Vi borjar med en precis definition av deriverbarhet :

DEFINITION : Lat $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^m$ dar D ar en oppen mangd, och lat $a \in D$. Vi sager att f ar *deriverbar* vid $x = a$ om det existerar en $m \times n$ matris A sa att, i narheten av a,

$$f(x) = f(a) + A(x-a) + \tau(x), \tag{31}$$

dar $\left|\frac{\tau(x)}{x-a}\right| \to 0$ nar $x \to a$.

Vi ska komma tillbaka till vektorvarda funktioner i kapitel 3, men genom resten av kapitel 2 vi ska bearbeta bara reellvarda funktioner. Sa antag att f ar en deriverbar, reellvard funktion i n variabler. Da matrisen A i (31) ar en matris med en enda rad, sag $A = (A_1, ..., A_n)$. Vi betecknar vektoren $x - a$ som $h = (h_1, ..., h_n)^T$. Da ar produkten $A(x-a)$ en skalar produkt och vi far formulen

$$f(x) = f(a) + \sum_{i=1}^{n} A_i h_i + \tau(x). \tag{32}$$

Antag nu att x gar mot a parallell med x_j-axeln, sa att $h_i = 0$ for alla $i \neq j$ och $h_j \to 0$. Vid gransen (in the limit), (2) blir

$$A_j = \lim_{h_j \to 0} \frac{f(a + h_j \mathbf{e}_j) - f(a)}{h_j}. \tag{33}$$

Markera att gransvardet pa hogra sidan ar derivaten av en 1-variabel funktion, namligen funktionen f inskrankt till den ratta linjen genom a parallell med x_j-axeln. Detta gransvarde kallas for den *partiella derivaten* av f vid *a med avseende pa* x_j, och betecknas $f_{x_j}(a)$. Sa har vi

$$f_{x_j}(a) = \lim_{t \to 0} \frac{f(a + t\mathbf{e}_j) - f(a)}{t}. \tag{34}$$

Fordelen av partiella derivator ar att de kan beraknas med metoder fran envariabelanalys. For att berakna den j:te partiella derivaten av f, man betraktar f som en funktion av den enda variabeln x_j med alla andra x_i

betraktas som konstanter.

Exempel 1 : Lat $f(x,y,z) = xyz^3 + \cos(xz) + e^{x^2+yz}$. Berakna alla partiella derivator vid $(1,1,1)$.

Losning : Vi far foljande allmana formuler for de partiella derivatorna, och da satter in $(1,1,1)$:

$$f_x = yz^3 - z\sin(xz) + 2xe^{x^2+yz},$$
$$f_y = xz^3 + ze^{x^2+yz},$$
$$f_z = 3xyz^2 - x\sin(xz) + ye^{x^2+yz}.$$

Fran foregaende diskussion ser vi att, om funktionen f ar deriverbar, da existerar alla dess partiella derivator. Livet vore mycket skont om motsatsen var sant - tyvarr, det finns icke-deriverbara funktioner vilkas partiella derivator existerar (exempel 2 nere). Men om de partiella derivatorna uppfyller ett svagt (weak) villkor, da ar deriverbarhet garanterad (sats 1 nere). Vi borjar med en definition :

Definition : Lat D vara en oppen mangd i $\mathbf{R}^n$. *Klassen* $C^1(D)$ ar definierad som mangden av alla reellvarda funktioner pa D vilkas partiella derivator existerar overallt i D och ar kontinuerliga funktioner (i D).

Nu har vi

Sats 1. *$f \in C^1(D) \Rightarrow f$ deriverbar overallt i D.*

Bevis : Jag hoppade over detta. Ideen ar att anvanda *medelvardssatsen* fran envariabelanalys. Beviset i boken ar lite svart att lasa, ty det finns sa manga olika beteckningar. Jag ska skriva upp mitt eget bevis. Titta ocksa i Adams' bok eller nagon annan amerikansk bok.

Nu vill vi ge ett exempel av en icke-deriverbar funktion vilkens partiella derivator existerar. Forst, en latt proposition :

Proposition 2. *f deriverbar vid $x = a \Rightarrow f$ kontinuerlig vid $x = a$.*

Bevis : Latt. Anvand definitionen av deriverbarhet.

EXEMPEL 2 : Betrakta funktionen

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases} \quad (35)$$

Med hjalp av (34), kan vi visa att $f_x(0,0) = f_y(0,0) = 0$. For en godtyclig punkt $(x,y) \neq (0,0)$ vi anvander kvot regeln mm for att erhalla

$$f_x = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2} \quad (36)$$

$$f_y = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}. \quad (37)$$

Darfor, existerar bada partiella derivator overallt i $\mathbf{R}^2$. Men det ar 'latt' att bevisa att

(a) f ar inte kontinuerlig vid $(0,0)$, alltsa inte deriverbar vid $(0,0)$.
(b) $f \notin C^1(\mathbf{R}^2)$ ty varken f_x eller f_y har ett gransvarde som $(x,y) \to (0,0)$.

OBS ! Det **finns** deriverbara funktioner som inte ar C^1 men, aven i det 1-dimensionella fallet, de ar inte sa latt att skapa. Denna ar en ovning pa den forsta inlamningsuppgiften !

$n = 2$: **Tangentplan**

Grafen $z = f(x,y)$ av en 2-variabel funktion ar en yta i rummet. Geometriskt, f ar deriverbar vid en punkt (x_0, y_0) omm det finns ett unikt plan tangent till denna yta vid $(x_0, y_0, z_0 = f(x_0, y_0))$. Fran definition (32) av deriverbarhet, far vi att planets ekvation ar

$$[x - x_0] f_x(x_0, y_0) + [y - y_0] f_y(x_0, y_0) - [z - z_0] = 0. \quad (38)$$

Om vi introducerar funktionen $g(x,y,z) := f(x,y) - z$ sa att grafen blir den implicita funktionen $g = 0$ av 3 variabler, da kan vi skriva planets ekvation som

$$\vec{\nabla} g(v_0) \ \ (v - v_0) = 0, \quad (39)$$

dar $\vec{\nabla} g := (g_x, g_y, g_x)$, $v := (x,y,z)$ och $v_0 := (x_0, y_0, z_0)$. Detta motiverar foljande definition

DEFINITION : Lat f vara en reellvard funktion av n variabler vilkas partiella derivator existerar. Den vektorvarda funktion

$$\vec{\nabla} f := (f_{x_1}, f_{x_2}, ..., f_{x_n}) \tag{40}$$

kallas for *gradienten av* f.

Egenskap/tillampningar av den gradient funktionen

Ideen ar att gradienten kan betraktas som en generalisering av en derivata i envariabelanalys. Detaljer nasta gang ....

4 MAN 030 : Sammandrag av föreläsning 4 Tisdag 14/9/99

Vi fortsätter med vår diskussion om den *gradienta funktionen.*

OBS ! f är $C^1 \Leftrightarrow \vec{\nabla} f$ är kontinuerlig.

Den gradienta funktionen är användbar -
1) för att förenkla beteckningar
2) för vidare teoretiska undersökningar (t.ex. i vektoranalys, som vi kommer att studera i del 2)
3) i tillämpningar.

Den allmäna idéen är att man kan upfatta gradienten som en korrekt generalisering av den derivaten av en vanlig en-variabel funktion.

Exempel 1 : Villkoret för deriverbarhet kan skrivas som

$$f(\vec{x}) = f(\vec{a}) + \vec{\nabla} f \cdot (\vec{x} - \vec{a}) + \tau(\vec{x}), \tag{41}$$

där $\left|\frac{\tau(\vec{x})}{\vec{x}-\vec{a}}\right| \to 0$ när $\vec{x} \to \vec{a}$, som vanligt.

Exempel 2 : I envariabelanalys, har vi följande faktum

Sats. $f' \equiv 0 \Rightarrow f = $ *konstant.*

Bevis : En lätt tillämpning av *medelvärdssatsen.*

I flervariabelanalys, finns det ingen meningsfull generalisering av medelvärdssatsen, men vi kan generalisera övanstående hävdande (assertion). Först, behöver vi ett par definitioner :

Definition 1 : En *båge* i $\mathbf{R}^n$ är en kontinuerlig funktion $\phi : [0,1] \to \mathbf{R}^n$. De punkter $a = \phi(0)$ och $b = \phi(1)$ kallas för bågens *ändpunkter*, och vi säger att bågen *anslutar a till b* (connects/joins a to b).

Definition 2 : En delmängd $D \subseteq \mathbf{R}^n$ kallas för *bågvis sammanhängande* om, för varje två punkter $a, b \in D$, det finns en båge, **i mängden** D, som anslutar a till b.

Då har vi

Sats. *Låt $D \subseteq \mathbf{R}^n$ vara öppen och bågvis sammanhängande, och $f : D \to \mathbf{R}$ deriverbar. Om $\vec{\nabla} f \equiv 0$, då är f konstant.*

Bevis (Översikt) : Låt $a, b \in D$. Vi måste visa att $f(a) = f(b)$. Låt $\phi : [0,1] \to D$ vara en båge som anslutar a till b, dvs $\phi(0) = a$ och $\phi(1) = b$. Definiera $s \in [0,1]$ som

$$s = \sup\{t \in [0,1] : f(\phi(u)) = f(a) \text{ för alla } u \in [0,t]\}. \tag{42}$$

Ty f är kontinuerlig, $f(\phi(s)) = f(a)$. Ty D är öppen och $\vec{\nabla} f \equiv 0$, ser vi att $s = 1$, dvs att $f(a) = f(b)$. q.e.d.

(Obs ! Detta bevis kan förklaras mycket bättre med hjälp av en tavla, där man kan rita bilder).

Exempel 3 (** viktigt ! **) : De partiella derivator av en funktion f vid en punkt beskriver 'the rate of change' (? svenska) av f i vissa riktningar, nämligen parallell med en av de koordinata axlerna. Men, förstås, för att förstå f's uppförande, måste man veta hur f varierar i alla andra riktningar också. Vi får en skön formul nere med hjälp av den gradienta funktionen. Först, den lämpliga definitionen :

Definition : Låt $f : D \subseteq \mathbf{R}^n \to \mathbf{R}$ vara deriverbar, $\vec{a} \in D$ och $\hat{v}$ en enhetsvektor i $\mathbf{R}^n$. Den *riktningsderivaten* av f vid $\vec{a}$, *i riktningen av* $\hat{v}$, definieras (och betecknas) som

$$\frac{\partial f}{\partial \hat{v}} = f_{\hat{v}} := \lim_{t \to 0} \frac{f(\vec{a} + t\hat{v}) - f(\vec{a})}{t}. \tag{43}$$

Notera att om $\hat{v} = \vec{\mathbf{e}}_{x_i}$, då är $\frac{\partial f}{\partial \hat{v}} = \frac{\partial f}{\partial x_i}$. Nu har vi följande formulen för riktningsderivaten :

Sats.

$$\frac{\partial f}{\partial \hat{v}}(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot \hat{v}. \tag{44}$$

Bevis : Sätt $\vec{x} = \vec{a} + t\hat{v}$ i ekv. (1), och tag gränsvärdet när $t \to 0$ .....

Följdsats. *f växer snabbast (vid en punkt) i riktningen parallell med $\vec{\nabla} f$.*

Exempel 4 : Låt

$$f(x_1, ..., x_n) = C \tag{45}$$

vara en niva-hyperyta till en funktion $f \in C^1(\mathbf{R}^n)$. Denna är en $(n-1)$-dimensionell hyperyta S i $\mathbf{R}^n$. Låt p vara en punkt på S. Då finns det ett $(n-1)$-dimensionell hyperplan P tangent till hyperytan i p.

Påstående : $\vec{\nabla} f(p)$ är en normal vektor till P.

Bevis : Använd (1). Ty f är konstant på S, får vi att $\vec{\nabla} f(p) \cdot d\vec{x}(p) = 0$ för varje $d\vec{x}$ på ytan S. Det vill säga att $\vec{\nabla} f(p)$ är ortogonell med P (kom ihåg att en skalärprodukt av två vektorer i $\mathbf{R}^n$ är lika med 0 omm de är ortogonella).

OBS! Våra observationer i exempel 3 och 4 används mycket i naturvetenskapliga tillämpningar, där funktionen $\vec{\nabla} f$ har en klar mening. T. ex., i fysik, om V är ett energifalt, då är $\vec{\nabla} V$ det anslutet (associated) kraftfältet $\vec{F}$. Från exempel 3 ser vi att $\vec{F}$ beskriver den bågen som en partikel ska följa för att nå 'a state of minimal energy' med avseende på V. Detta förklarar termen 'kraftfält'.

Kedjeregeln

I envariabelanalys, finns det några elementära funktioner (t.ex. $x^n, \sin(x), e^x$) vars derivator kan beräknas direkt från definitionen. För att skapa mer komplicerade funktioner, tar vi summor, produkter, kvot och sammansättningar av sådana funktioner. Alltså, för att kunna differentiera dessa, hittar vi formuler för derivaten av en summa, produkt, kvot eller sammansättning. Alla dessa formuler kan generaliseras till högre dimensioner. Den viktigaste, och svåraste att behandla, är formulen för en sammansättning, som kallas för *kedjeregeln*.

Kom ihåg den kedjeregeln från 1-variabel analys : låt $f, g : \mathbf{R} \to \mathbf{R}$ vara deriverbara, och definiera den sammansätta funktionen $h = f \circ g$ (läs : 'g efter f') som $h(x) = f(g(x))$. Då får vi

Sats. *h är deriverbar och*

$$h'(a) = f'(g(a)) \cdot g'(a) \qquad (a \in \mathbf{R}). \tag{46}$$

Vi har följande generalisering till högre dimensioner :

Sats (kedjeregeln i Euklidiska rum). *Låt $g : \mathbf{R}^n \to \mathbf{R}^m$ och $f : \mathbf{R}^m \to \mathbf{R}$. Antag att båda f och g är C^1 och betrakta den sammansätta funktionen $h = f \circ g : \mathbf{R}^n \to \mathbf{R}$. Då är h också C^1 och*

$$\frac{\partial h}{\partial x_i}(\vec{a}) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(g(\vec{a})) \cdot \frac{\partial g_k}{\partial x_i}(\vec{a}) \qquad (\vec{a} \in \mathbf{R}^n,\ 1 \leq i \leq n), \tag{47}$$

där $g_k : \mathbf{R}^n \to \mathbf{R}$ är den k:te komponenta funktionen av g, så att $g(\vec{x}) = (g_1(\vec{x}), ..., g_m(\vec{x}))$.

Bevis : För långt att skriva ut här. Idéen är att använda faktum att en funktion ζ är deriverbar vid en punkt $\vec{a} \Leftrightarrow \zeta(\vec{x}) - \zeta(\vec{a})$ kan approximeras med en linjär funktion i närheten av $\vec{a} \Leftrightarrow$ det kan approximeras med $M(\vec{x} - \vec{a})$ för någon matris M. Om matriserna B, C svarar mot de deriverbara funktionerna g, f vid $\vec{a}, g(\vec{a})$ respektivt, vi kan visa att den produkta matrisen CB svarar mot h vid $\vec{a}$. Därifrån, får vi (47).

5 MAN 030 : Sammandrag av föreläsning 5 Torsdag 16/9/99

I envariabelanalys, kedjeregeln kan skrivas som

$$h'(a) = f'(g(a)) \cdot g'(a), \tag{48}$$

där $h = f \circ g$. Med några lämpliga definitioner, regeln kan skrivas på en liknande sätt i högre dimensioner.

Definition : Låt $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^m$ vara en C^1 funktion, och $\vec{a} \in D$. Den $m \times n$ matris $A = (a_{ij})$, där

$$a_{ij} = \frac{\partial f_i}{\partial x_j}(\vec{a}), \tag{49}$$

kallas för den *derivaten av* f vid $\vec{a}$, och betecknas $f'(\vec{a})$.

Nu kan den mest allmäna formuleringen av kedjegregeln skrivas som följer

Sats (Kedjeregeln). *Låt D, E vara öppna mängder i $\mathbf{R}^n$ och $\mathbf{R}^m$ resp. Låt $g : D \to E$ och $f : E \to \mathbf{R}^p$ vara C^1 funktioner. Sätt $h = f \circ g$. Då är h en C^1 funktion och, för varje $\vec{a} \in D$,*

$$h'(\vec{a}) = f'(g(\vec{a})) \cdot g'(\vec{a}), \tag{50}$$

där $\cdot$ betecknar multiplikation av matriser.

Bevis : Ty f och g är differentierbara får vi, med hjälp av definitionen av differentierbarhet, att, för $\vec{x}$ tillräckligt nära $\vec{a}$,

$$h(\vec{x}) - h(\vec{a})$$

$$= f(g(\vec{x})) - f(g(\vec{a}))$$

$$= f'(g(\vec{a})) \cdot (g(\vec{x}) - g(\vec{a})) + \tau_1(g(\vec{x})), \quad \text{där } \left|\frac{\tau_1(g(\vec{x}))}{g(\vec{x}) - g(\vec{a})}\right| \to 0, \text{ när } \vec{x} \to \vec{a}, \quad ,$$

$$= f'(g(\vec{a})) \cdot [g'(\vec{a}) \cdot (\vec{x} - \vec{a}) + \tau_2(\vec{x})] + \tau_1(g(\vec{x})), \quad \text{där } \left|\frac{\tau_2(\vec{x})}{\vec{x} - \vec{a}}\right| \to 0, \text{ när } \vec{x} \to \vec{a},$$

$$= f'(g(\vec{a})) \cdot g'(\vec{a}) \cdot (\vec{x} - \vec{a}) + \tau(\vec{x}),$$

där $\tau(\vec{x}) := \tau_1(g(\vec{x})) + f'(g(\vec{a})) \cdot \tau_2(\vec{x})$.

Det följer att h är differentierbar vid $\vec{a}$, med derivaten $f'(g(\vec{a})) \cdot g'(\vec{a})$ efter behov, ty $\left|\frac{\tau(\vec{x})}{\vec{x}-\vec{a}}\right| \to 0$ när $\vec{x} \to \vec{a}$ (som följer lätt från definitionen av τ_1 och τ_2).

Anmärkning : Den derivaten av en vektorvärd C^1-funktion f kallas också för den *funktionalmatris* till f. Denna terminologi används mer i funktionalanalys.

En viktig typ av vektorvärd funktion är de så kallade 'koordinat transformationer' -

Definition : En injektiv funktion $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ (samma n !) kallas för en *(n-dimensionell) koordinat transformation.*

Vi är interesserad av koordinat transformationer som är också, åtminstone, C^1-funktioner. Notera att definitionsmängden av en koordinat transformation (dvs, den storste mängden där funktionen ar injektiv) är ofta inte en öppen mängd i $\mathbf{R}^n$ (exempel 1, nere). Om vi säger att en sådan funktion är C^1, menar vi att den är C^1 i varje öppen delmängd av dess definitionsmängd. Man kan visa, faktiskt, att en union av öppna mängder är öppen (GLO, s. 18 ff) så att varje delmängd D av $\mathbf{R}^n$ har en unik storste öppen delmängd, som kallas för *inre* av D, och betecknas $\mathrm{Int}(D)$.

Definition : Låt $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ vara en C^1 koordinat transformation. Låt $\vec{a} \in \mathrm{Int}(D)$. Den derivaten av f vid $\vec{a}$ kallas också för den *Jacobian matrisen* till f, och betecknas $\mathrm{Jac}[f](\vec{a})$.

Exempel 1 : Den transformation från rättvinkliga till sfäriska koordinator i $\mathbf{R}^3$ skrivas som

$$F(r, \theta, \phi) = (x, y, z) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta). \tag{51}$$

Den definitionsmängd D till F är

$$D = (0, +\infty) \times [0, \pi) \times [0, 2\pi), \tag{52}$$

så att $\mathrm{Int}(D) = (0, +\infty) \times (0, \pi) \times (0, 2\pi)$. Den Jacobian till F vid en godtycklig punkt i $\mathrm{Int}(D)$ är

$$\mathrm{Jac}[f] = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}. \tag{53}$$

6 Sammandrag av Föreläsning 6 Tisdag 21/9/99

Invers funktion satsen

Motivering från envariabelanalys :

Låt $f : \mathbf{R} \to \mathbf{R}$ vara C^1 och antag att $f'(a) \neq 0$. Vi gör följande observationer :

1. Ty f' är kontinuerlig, det finns $\delta > 0$ så att $f'(x) \neq 0$ för alla x i det öppna intervallet $I = (a - \delta, a + \delta)$.

2. Då Rolle's sats (ett speciellt fall av den Medelvärdssatsen) medförer att f är injektiv på I.

3. Den Mellanliggandevärdssatsen (?) medförer, faktiskt, att f är antingen växande eller avtagande på I.

4. Då är det klart att $f(I)$ är ett öppet intervall $J \ni f(a)$.

5. Ty $f|_I$ är injektiv, det existerar en invers funktion g på $f(I)$. Ty $f(I)$ är ett öppet intervall, vi kan fråga om g är också C^1. Kedjeregeln antyder att g ÄR C^1 och att

$$g'(f(x)) = \frac{1}{f'(x))} \tag{54}$$

för alla $x \in I$.

OBS! f kan vara injektiv i närheten av a, även om $f'(a) = 0$. Men i detta fall, den invers funktionen skall inte vara differntierbar vid a.

Exempel : Betrakta $f(x) = x^3$. f är en C^1 bijektion från $\mathbf{R}$ till $\mathbf{R}$, med invers $g(x) = \sqrt[3]{x}$. Notera att g är inte differentierbar vid $x = 0$, som svarar mot faktum att $f'(0) = 0$.

Vi har följande generalisering av övanstående observationer till högre dimensioner :

Sats (Invers funktion sats) : *Låt $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ vara en C^1 funktion* *och låt $\vec{a} \in D$ vara en punkt så att*

$$det(Jac[f](\vec{a})) \neq 0. \tag{55}$$

Då existerar det en öppen omgivning U av $\vec{a}$ så att $f|_U$ är injektiv och $f(U)$ är en öppen omgivning V av $f(\vec{a})$. Den invers funktionen $g : V \to U$ är också C^1 och, för alla $\vec{x} \in U$, får vi att

$$Jac[g](f(\vec{x})) \cdot Jac[f](\vec{x}) = E_n, \tag{56}$$

där E_n är den $n \times n$ enhets matrisen.

Exempel : Nr. 3.36 i övningsboken.

Impicit funktion satsen

Motivering :

Låt $F(x,y)$ vara en reellvärd funktion av två variabler och betrakta en nivå kurva

$$F(x,y) = C. \tag{57}$$

Vi vet att (57) inte alltid representerar en vanlig 1-variabel C^1 funktion $y = f(x)$. T.ex. om $F(x,y) = x^2 + y^2$, varje nivå kurva (för $C > 0$) är en cirkel. Men även cirkeln kan beskrivas LOKALT i formen $y = f(x)$, med undantag av två punkter där tangenten är vertikal, dvs där $\frac{\partial F}{\partial y} = 0$. På samma sätt, cirkeln kan också beskrivas lokalt i formen $x = f(y)$ med undantag av två punkter där $\frac{\partial F}{\partial x} = 0$.

Definition : Låt $F(x,y)$ vara en C^1 reellvärd funktion. Låt (a,b) vara en punkt på nivåkurvan (57). Vi säger att (57) definierar y *implicit* (eng.: implicitly defines y) som en funktion av x, *i närheten av* (a,b), om det existerar ett öppet intervall $I \ni a$ så att, för varje $x \in I$, det finns en unik $y = f(x)$ så att (x,y) ligger på kurvan (57).

Då har vi följande generalisering av övanstående observationer :

Sats (Implicit funktion sats i 2-d) : *Låt $F(x,y)$ vara en C^1 reellvärd funktion och betrakta en nivåkurva (57). Låt (a,b) vara en punkt på kurvan.*

Då (57) definierar y implicit som en funktion av x, i närheten av (a, b), om och endast om

$$\frac{\partial F}{\partial y}(a, b) \neq 0. \tag{58}$$

OBS! Om (58) stämmer och vi skriver den implicita definierade funktionen i formen $y = f(x)$, då den metoden av *implicit derivering* antyder att

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y}. \tag{59}$$

BEVIS AV SATSEN : Man tillämpar den Invers funktion satsen till funktionen

$$G(x, y) := (x, F(x, y)). \tag{60}$$

Detaljer gavs i föreläsningen. Notera att inget bevis ges i boken, men ni behöver mitt (eller något annat) bevis för tentamen.

Det finns en liknande version av den Implicit funktion satsen, med samma bevis, i varje $\mathbf{R}^n$, nämligen -

Sats (Implicit funktion sats i n-d) : *Låt $F(x_1, ..., x_n)$ vara en reellvärd funktion av n reella variabler, och betrakta en nivåhyperyta S med ekvation $F(x_1, ..., x_n) = C$. Låt $\vec{a} \in S$. Då definieras x_n implicit som en funktion av $x_1, ..., x_{n-1}$, i närheten av $\vec{a}$ på S, om och endast om $\frac{\partial F}{\partial x_n}(\vec{a}) \neq 0$.*

OBS! Den mest allmäna versionen av den implcit funktion satsen ges i boken på sidan 134.

Det finns följande tillämpning av den Implicit funktion satsen till ytor i $\mathbf{R}^3$:

Sats. *Låt $F(x, y, z)$ och $G(x, y, z)$ vara två reellvärda C^1-funktioner av tre variabler, och betrakta de nivåytor S_1, S_2 med ekvationer*

$$F(x, y, z) = C$$
$$G(x, y, z) = D$$

respektivt. Låt $\vec{a} \in S_1 \cap S_2$. Då finns det en C^1-kurva γ genom $\vec{a}$, som ligger helt i $S_1 \cap S_2$, om och endast om

$$[\vec{\nabla} F \times \vec{\nabla} G](\vec{a}) \neq \vec{0}. \tag{61}$$

Dessutom, om den x*- (resp.* y*-* z*-) komponenten av denna vektorprodukt är inte lika med noll, då kan* γ *paramateriseras med* x *(resp.* y z*) som parameter.*

Bevis : Behövs inte, men samma idé som i beviset av den Implicit funktion satsen. En intuitiv, geometrisk tolkning av denna sats gavs i föreläsningen.

7 Sammandrag av föreläsning 7 Torsdag 23/9/99

Partiella derivator av ordning ≥ 2

DEFINITION : Låt D vara en öppen delmängd av $\mathbf{R}^n$. Vi definierar

$$C^0(D) := \{f : D \to \mathbf{R} : f \text{ är kontinuerlig}\} \tag{62}$$

och

$$C^1(D) := \{f \in C^0(D) : \text{ alla } \tfrac{\partial f}{\partial x_i} \text{ existerar överallt i } D \text{ och är kontinuerliga}\}. \tag{63}$$

Induktivt, definierar vi, för $k > 1$,

$$C^k(D) := \{f \in C^1(D) : \text{ alla } \tfrac{\partial f}{\partial x_i} \text{ ligger i } C^{k-1}(D)\}, \tag{64}$$

och, äntligen,

$$C^\infty(D) = \bigcap_{k \geq 0} C^k(D). \tag{65}$$

Notera att $f \in C^k(D)$ om och endast om alla partiella derivator av f, av ordning $\leq k$, existerar och är kontinuerliga överallt i D. Och $f \in C^\infty(D)$ omm alla partiella derivator av alla ordningar existerar och är kontinuerliga överallt i D. Ett element av $C^\infty(D)$ kallas också för en *jämn (eng.: smooth)* funktion på D.

Nästa sats betyder att ordningen spelar ingen roll när vi vill beräkna en blandad partiell derivata av någon ordning ≥ 2.

Sats. *Låt $f : \mathbf{R}^n \to \mathbf{R}$ vara C^2. Då är*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} \quad \textit{för alla } i, j. \tag{66}$$

BEVIS : Det räcker att bevisa satsen när $n = 2$. Då följade jag beviset i boken, som är tekniskt, men detta verkar vara någonting som inte kan undvikas.

Taylor's sats

NOTATION 1 : Låt f, g vara två reellvärda funktioner definierad i mängden S. Vi skriver

$$f(x) = O(g(x)) \tag{67}$$

om det finns en konstant $K > 0$ så att $|f(x)| \leq K|g(x)|$ för alla $x \in S$.

Nu, kom ihåg Taylor's sats i $\mathbf{R}$:

Sats. *Låt $f : D \subseteq \mathbf{R} \to \mathbf{R}$ vara C^{k+1} och låt $a \in D$. Då existerar det en omgivning $I \subseteq D$ av a, så att*

$$f(x) = f(a) + \sum_{i=1}^{k} \frac{f^{(i)}(a)}{i!}(x-a)^i + O[(x-a)^{k+1}], \tag{68}$$

för alla $x \in I$.

För att kunna skriva ner generaliseringen till högre dimensioner, behöver vi mer tekniska notationer :

NOTATION 2 : I $\mathbf{R}^n$ införer vi den vektorvärda *operatoren*

$$\vec{\nabla} := \left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}\right) \tag{69}$$

Låt $\vec{h} = (h_1, ..., h_n) \in \mathbf{R}^n$. Vi sätter

$$\vec{h} \cdot \vec{\nabla} := \sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} \tag{70}$$

och, för $m \geq 1$,

$$(\vec{h} \cdot \vec{\nabla})^m = (\sum_i h_i \frac{\partial}{\partial x_i})^m := \sum_{\substack{\text{alla } (k_1,...,k_n) \text{ så att} \\ k_1+...+k_n=m}} \left[\prod_{i=1}^{n} h_i^{k_i} \frac{\partial^m}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}\right]. \tag{71}$$

För en C^m-funktion $f : \mathbf{R}^n \to \mathbf{R}$, den operatoren $(\vec{h} \cdot \vec{\nabla})^m$ *opererar* på f genom formulen

$$(\vec{h} \cdot \vec{\nabla})^m \mathbf{f} = (\sum_i \frac{\partial}{\partial x_i})^m \mathbf{f} := \sum_{\substack{\text{alla } (k_1,...,k_n) \text{ så att} \\ k_1+...+k_n=m}} \left[\prod_{i=1}^{n} h_i^{k_i} \frac{\partial^m \mathbf{f}}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}\right]. \tag{72}$$

Notera att summan på högra sidan av (72) är en summa över alla möjliga partiella derivator av ordning m, med koefficienter som är *monom* (eng.: monomials) av grad m i de h_i.

Nu får vi

Taylor's sats i $\mathbf{R}^n$: *Låt $f : D \subseteq \mathbf{R}^n \to \mathbf{R}$ vara en C^{k+1} funktion, och låt $\vec{a} \in D$. Då existerar det en omgivning $U \subseteq D$ av $\vec{a}$ så att*

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \sum_{i=1}^{k} \frac{(\vec{h}\cdot\vec{\nabla})^i f}{i!} + O(|\vec{h}|^{k+1}), \tag{73}$$

för alla $\vec{a}+\vec{h} = \vec{x} \in U$.

BEVIS : Behövs inte, men idéen är att reducera till det 1-dimensionella fallet.

SPECIELLT FALL : $n = 2, k = 2$.

Bara det här fallet diskuteras i boken. Om vi låter $\vec{a} = (a,b)$ och $\vec{h} = (h,k)$, då (73) blir

$$\begin{aligned} f(a+h,b+k) = f(a,b) + \left[h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b)\right] \\ + \left[h^2\frac{\partial^2 f}{\partial x^2}(a,b) + 2hk\frac{\partial^2 f}{\partial x \partial y}(a,b) + k^2\frac{\partial^2 f}{\partial y^2}(a,b)\right] + O((h^2+k^2)^{3/2}), \end{aligned} \tag{74}$$

för alla (h,k) i en omgivning U av $(0,0)$.

LITE MER ALLMÄN : $k = 2$, n godtycklig.

Då (73) blir

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \left[\sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\vec{a})\right] + \left[\sum_{i=1}^{n} h_i^2 \frac{\partial^2 f}{\partial x_i^2}(\vec{a}) + \sum_{i\neq j} 2h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})\right] + O(|\vec{h}|^3), \tag{75}$$

för alla $\vec{h}$ i en omgivning U av origon.

Tillämpning till lokala extrema

Definition : (i) Låt $f : \mathbf{R}^n \to \mathbf{R}$ vara en (godtycklig !) funktion. Låt $\vec{a} \in \mathbf{R}^n$. Vi säger att f har en *lokal maximum* i punkten $\vec{a}$ om det finns en omgivning U av $\vec{a}$ så att $f(\vec{x}) \leq f(\vec{a})$ för alla $\vec{x} \in U$.

(ii) Samma definition för en *lokal minimum.*

(iii) Vi säger att f har en *lokal extremum* i punkten $\vec{a}$ om den har antingen en lokal max. eller en lokal min. vid $\vec{a}$.

Proposition. *Om $f : \mathbf{R}^n \to \mathbf{R}$ har en lokal extremum vid $\vec{a}$, då är*

$$\frac{\partial f}{\partial x_i}(\vec{a}) = 0, \qquad (i = 1, ..., n). \tag{76}$$

Bevis : Längs linjen genom $\vec{a}$ parallell med den x_i-axeln, f kan betraktas som en funktion av den enda variabeln x_i som har en extrem punkt vid $\vec{a}$. Då får vi (76) från envariabel analys.

Definition : En punkt där (76) satisfieras kallas för en *stationär punkt* eller *kritisk punkt* till f.

OBS! Från envariabel analys, vet vi att en stationär punkt är inte alltid en lokal extremum, t.ex.: $y = x^3$ vid $x = 0$.

I 1-dimension, kom ihåg att följande sats ger ett TILLRÄCKLIGT villkor för existens av en lokal extremum :

Sats. *Låt $f \in C^3(\mathbf{R})$ och antag att f har en stationär punkt vid $x = a$. Då är denna punkt en lokal extremum om och endast om $f''(a) \neq 0$. Vi får en lokal max. omm $f''(a) < 0$ och en lokal min. omm $f''(a) > 0$.*

I högre dimensioner, finns det också ett tillräckligt villkor som uttryckas i termer av de värden av de partiella derivator av f av ordning två i den stationära punkten. Men villkoret är mycket svårare att beskriva, så vi måste ta det långsamt !

Först, notera att om f är C^3 och har en stationär punkt i $\vec{a}$, då (75) säger att

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \left[\sum_{i=1}^{n} h_i^2 \frac{\partial^2 f}{\partial x_i^2}(\vec{a}) + \sum_{i\neq j} 2h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})\right] + O(|\vec{h}|^3), \quad (77)$$

för alla tillräckliga små $\vec{h}$. Sätt

$$a_{ii} := \frac{\partial^2 f}{\partial x_i^2}(\vec{a}) \quad a_{ij} := \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) \ \ (i \neq j) \quad (78)$$

$$Q(\vec{h}) = Q(h_1, ..., h_n) := \sum_{i=1}^{n} a_{ii} h_i^2 + \sum_{i\neq j} 2a_{ij} h_i h_j, \quad (79)$$

så att (77) blir

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + Q(\vec{h}) + O(|\vec{h}|^3). \quad (80)$$

Intuitivt, är det klart att f's uppförande nära $\vec{a}$ bestämms av funktionen Q. Denna är en *homogen* kvadratisk funktion av n variabler, som kallas för en *kvadratisk form* i n variabler. Alltså, innan vi kan ställa fram vårt tillräckligt villkor, måste vi studera kvadratiska former.

Viktig anmärkning : I tillämpningar, man är ofta interesserad av att hitta *globala*, i stället för lokala, extrema av en funktion. Med en kombination av den lokala teorin och den så kallade *Heine-Borel satsen* (GLO, kap. 1), det globala problemet kan ofta lösas, särskilt för funktioner med *kompakta* (GLO, kap. 1) definitionsmängder. To be continued ! ...

8 Sammandrag av föreläsning 8 Tisdag 28/9/99

Vi börjar med två olika definitioner av en kvadratisk form :

DEFINITION 1 : En *kvadratisk form* i de n reella variablerna $x_1, ..., x_n$ är en funktion $Q : \mathbf{R}^n \to \mathbf{R}$ av formen

$$Q(x_1, ..., x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{i \neq j} 2a_{ij} x_i x_j, \tag{81}$$

där a_{ii}, a_{ij} är reella tal.

DEFINITION 2 : En *kvadratisk form* av n reella variabler är en funktion $Q : \mathbf{R}^n \to \mathbf{R}$ så att

$$Q(\vec{0}) = \vec{0}, \quad Q(\alpha\vec{x}) = \alpha^2 Q(\vec{x}) \text{ för alla } \alpha \in \mathbf{R},\, \vec{x} \in \mathbf{R}^n. \tag{82}$$

Det är lätt att se att (1) $\Rightarrow$ (2), men inte helt triviallt i den andra riktningen.

Vi delar upp kvadratiska former i fem olika klasser :

DEFINITION 3 : (i) En kvadratisk form kallas för *positiv definit* om $Q(\vec{x}) > 0$ för alla $\vec{x} \neq \vec{0}$.

(ii) Q kallas för *negativ definit* om $Q(\vec{x}) < 0$ för alla $\vec{x} \neq \vec{0}$.

(iii) Q kallas för *positiv semi-definit* om $Q(\vec{x}) \geq 0$ för alla $\vec{x}$ och det finns en $\vec{x_0} \neq \vec{0}$ så att $Q(\vec{x_0}) = 0$.

(iv) Q kallas för *negativ semi-definit* om $Q(\vec{x}) \leq 0$ för alla $\vec{x}$ och det finns en $\vec{x_0} \neq \vec{0}$ så att $Q(\vec{x_0}) = 0$.

(v) Q kallas för *indefinit* annars, dvs om det finns $\vec{x_1}, \vec{x_2}$ så att $Q(\vec{x_1}) > 0$ och $Q(\vec{x_2}) < 0$.

Vårt viktigaste mål är att beskriva en sats som gör det lätt att bestämma sig till vilken klass hör en godtycklig form Q. Först, behöver vi en definiton till :

DEFINITION 4 : En koordinat transformation $\phi : \mathbf{R}^n \to \mathbf{R}^n$ kallas för *ortogonal* om

$$|\phi(\vec{x})| = |\vec{x}| \quad \text{för alla } \vec{x} \in \mathbf{R}^n. \tag{83}$$

Sats (Sylvester's sats/ Witt's sats) : *Låt $Q : \mathbf{R}^n \to \mathbf{R}$ vara en kvadratisk form i de n variablerna $x_1, ..., x_n$. Då finns det en ortogonal koordinat transformation $(x_1, ..., x_n) \xrightarrow{\phi} (y_1, ..., y_n)$ så att, som en funktion av de nya variablerna $y_1, ..., y_n$, Q kan uttryckas som*

$$Q(y_1, ..., y_n) = \sum_{i=1}^{n} \lambda_1 y_i^2 \quad (\lambda_i \in \mathbf{R}). \tag{84}$$

Dessutom, de koefficienterna λ_i är oberoende av valet av ortogonal transformation.

Bevis : Linjär algebra. Behövs inte.

Anmärkning : En kvadrarisk form av typ (84) kallas för *diagonal.* De koefficienterna λ_i kallas för formens *egenvärde.*

Nästa proposition följer direkt från definitonerna :

Proposition. *Låt Q vara en diagonal kvadratisk form av typ (84). Då*

(i) Q positiv definit $\Leftrightarrow$ alla $\lambda_i > 0$.
(ii) Q negativ definit $\Leftrightarrow$ alla $\lambda_i < 0$.
(iii) Q positiv semi-definit $\Leftrightarrow$ alla $\lambda_i \geq 0$ och någon $\lambda_i = 0$.
(iv) Q negativ definit $\Leftrightarrow$ alla $\lambda_i \leq 0$ och någon $\lambda_i = 0$.
(v) Q indefinit $\Leftrightarrow$ någon $\lambda_i > 0$ och någon annan $\lambda_i < 0$.

Nästa sats ger tillämpningen till klassificeringen av stationära punkter. Först, en definition :

Definition : En stationär punkt som är varken en lokal max. eller en lokal min. kallas för en *sadelpunkt.*

Sats 2 **. *Låt $f \in C^3(\mathbf{R}^n)$. Antag att f har en stationär punkt i $\vec{a}$ och en Taylor utveckling*

$$f(\vec{a} + \vec{h}) = f(\vec{a}) + Q(h_1, ..., h_n) + O(|\vec{h}|^3), \tag{85}$$

där

$$Q(h_1, ..., h_n) = \sum_{i=1}^{n} a_{ii} h_i^2 + \sum_{i \neq j} 2a_{ij} h_i h_j \tag{86}$$

och

$$a_{ii} = \frac{\partial^2 f}{\partial x_i^2}(\vec{a}), \quad a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{87}$$

Då,

(i) Om Q är positiv definit $\Rightarrow$ f har en lokal minimum i $\vec{a}$.
(ii) Om Q är negativ definit $\Rightarrow$ f har en lokal maximum i $\vec{a}$.
(iii) Om Q är indefinit $\Rightarrow$ f har en sadelpunkt i $\vec{a}$.

OBS! Om Q är semi-definit, vi kan inte dra någon slutsats om typen av den stationära punkten $\vec{a}$. Man måste använder någon annan metod för att försöka klassificera den. Tyvärr, det finns ingen kännt metod som fungerar i allmänhet.

Bevis av satsen : Del (iii) är klart. Bevis av del (i) och del (ii) gavs i föreläsningen, och dess viktigast ingrediens är följande lemma :

Lemma : Om en kvadratisk form Q är positv definit (resp. negativ definit), då finns det en konstant $K > 0$ så att $Q(\vec{x}) > K|\vec{x}|$ (resp. $Q(\vec{x}) < -K|\vec{x}|$) för alla $\vec{x} \neq \vec{0}$.

Bevis av lemma : Från Sylvester's sats ser vi att det räcker att bevisa lemman för diagonala former, och i det här fallet, det är ganska lätt. Mer detaljer gavs i föreläsningen.

Låga dimensioner : n = 2

För att sats 2 skulle vara användbar, vi måste kunna diagonalisera en godtycklig kvadratisk form. Det här är ett linjär algebraiskt problem, som vi skall inte försöka formulera i allmänhet. Men för $n = 2$ får vi en mycket lätt och explicit formulering av sats 2, nämligen

Sats 2′. *Låt $f(x, y)$ vara en C^3 funktion. Antag att (a, b) är en stationär punkt. Sätt*

$$A = f_{xx}(a, b) \quad B = f_{xy}(a, b) \quad C = f_{yy}(a, b). \tag{88}$$

Då,

(i) Om $AC - B^2 > 0$ och $A > 0 \Rightarrow f$ har en lokal minimum i (a,b).
(ii) Om $AC - B^2 > 0$ och $A < 0 \Rightarrow f$ har en lokal maximum i (a,b).
(iii) Om $AC - B^2 < 0 \Rightarrow f$ har en sadelpunkt i (a,b).

ANMÄRKNING : Om $AC-B^2 = 0$ (som svarar mot en semi-definit kvadratisk form), då kan vi inte dra någon slutsats om typen av den stationära punkten (a,b).

EXEMPEL : Hitta och klassificera alla stationära punkter till $f(x,y) = 3xy - x^3 - y^3$.

GEOMETRISK UPPFATTNING FÖR n = 2

Betrakrta en godtycklig 2-variabel kvadratisk form

$$Q(x,y) = Ax^2 + 2Bxy + Cy^2. \tag{89}$$

Låt grafen $z = Q(x,y)$ kallas för G. Då får vi att

(i) Q positiv definit $\Leftrightarrow AC - B^2 > 0$ och $A > 0 \Leftrightarrow G$ är en skål.
(ii) Q negativ definit $\Leftrightarrow AC - B^2 > 0$ och $A < 0 \Leftrightarrow G$ är en skål upp och ner.
(iii) Q indefinit $\Leftrightarrow AC - B^2 < 0 \Leftrightarrow G$ är en sadel.
(iv) Q semi-definit $\Leftrightarrow AC - B^2 = 0 \Leftrightarrow G$ är en dubbelt oändlig halvcylinder.

9 Sammandrag av föreläsning 9 Torsdag 30/9/99

Definition av de reella talen

Definition 1 : Låt Y vara en godtycklig mängd. En *följd* i Y är en funktion $f : \mathbf{N} \to Y$.

Om $f(k) = y_k$ vi brukar beteckna följden som $(y_k)_{k=1}^{\infty}$ eller, kort och gott, (y_k).

Definition 2 : Låt q vara ett rationellt tal. Vi sätter

$$|q| :\overset{\text{def}}{=} \begin{cases} q, \text{ om } q \geq 0. \\ -q, \text{ om } q < 0. \end{cases} \tag{90}$$

Definition 3 : Låt (q_k) vara en följd av rationella tal. Följden kallas för *Cauchy* om, för varje $n > 0$, det finns $N >> 0$ så att

$$|q_k - q_l| < \frac{1}{n}, \quad \text{för alla } k, l > N. \tag{91}$$

Definition 4 : Låt (q_k) och (p_k) vara två Cauchy följder av rationella tal. Vi säger att de två följder är *ekvivalenta* om, för varje $n > 0$, det finns $N >> 0$ så att

$$|q_k - p_l| < \frac{1}{n}, \quad \text{för alla } k, l > N. \tag{92}$$

Definition 5 : Ett *reellt tal* är en klass av ekvivalenta Cauchy följder av rationella tal.

Supremumaxiomet

I GLO boken, det supremumaxiomet för reella tal (som också kallas för det *fullständighetsaxiomet (eng.: completeness axiom)* presenteras som ett 'axiom', dvs någonting som inte kan bevisas och måste accepteras som given. Emellertid, med övanstående definition av de reella talen, supremumaxiomet blir en **sats**, dvs någonting som kan bevisas från mer fundamentala axiom.

Innan vi skriver ner satsen (som vi skall inte bevisa), behöver vi fler definitoner.

DEFINITION 6 : Låt X vara en delmångd av $\mathbf{R}$. Vi säger att X är *uppåt begränsad* (resp. *neråt begränsad*) om det finns ett reellt tal r så att $x \leq r$ (resp. $x \geq r$) för alla $x \in X$.

Talet r kallas för en *övre begränsning* (resp. *nedre begränsning*) till mången X.

DEFINITION 7 : Låt X vara en uppåt (resp. neråt) begränsad mångd. Det reella talet r_0 kallas för den *minsta övre begränsning* (resp. *högsta nedre begränsning*) till X om

(i) r_0 är en övre (resp. nedre) begränsning till X.

(ii) $r_0 \leq r$ (resp. $r_0 \geq r$) för alla andra övre (resp. nedre) begränsningar r till X.

DEFINITION 8 : En delmångd X av $\mathbf{R}$ kallas för *begränsad* om X är både uppåt och neråt begränsad.

ANMÄRKNING : En mer vanlig terminologi än 'minst övre (resp. högst nedre) begränsning' är *supremum* (resp. *infimum*), som brukar förkortas till *sup* (resp. *inf*) eller (på engelska) *lub* (resp. *glb*).

Sats 0 (Supremumaxiomet). *Varje uppåt begränsad mängd av reella tal har en minst övre begränsning.*

Följder av reella tal

Vi har redan infört begreppet av en Cauchy följd av rationella tal. Nu gör vi exakt samma för reella tal :

Definition 9 : Låt (r_k) vara en följd av reella tal. Följden kallas för *Cauchy* om, för varje $\epsilon > 0$, det finns $N >> 0$ så att

$$|r_k - r_l| < \epsilon, \quad \text{för alla } k, l > N. \tag{93}$$

Det finns ett annat intuitivt begrepp för följder av reella tal, nämligen konvergens. Den precisa definitionen av detta begrepp är :

Definition 9 : Låt (r_k) vara en följd av reella tal. Vi säger att följden *konvergerar* till det reella talet l om, för varje $\epsilon > 0$, det finns $N >> 0$ så att

$$k > N \Rightarrow |r_k - l| < \epsilon. \tag{94}$$

Den är ingen överraskning att de två begreppen i definitioner 8 och 9 är ekvivalenta. I en riktning, den här är lätt att bevisa. Vi har -

Proposition. *Varje konvergent följd av reella tal konvergerar.*

Bevis : Gavs i föreläsningen. Se också GLO.

Den sats som går i den andra riktningen är inte så lätt att bevisa - den beskriver en fundamental egenskap av de reella talen som kallas för *Cauchy's konvergens princip*. Den är ekvivalent med det Supremumaxiomet, och beviset beror av två andra satser, den *Monoton Konvergens satsen* och den *Bolzano-Weierstrass egenskapen*. Den första av dessa, MKS, är ekvivalent med Sup.ax., och är därför en särskilt egenskap av de reella talen. Den andra, BWE, gäller en godtycklig mångd med en ordnings relation '$\leq$'.

Innan vi skriver ner de här satserna, behöver vi en definition till :

Definition 10 : En följd (r_k) av reella tal kallas för *växande* (resp. *avtagande*) om $r_k \leq r_{k+1}$ (resp. $r_k \geq r_{k+1}$) för alla k.

En följd som är antingen växande eller avtagande kallas för *monoton*.

Anmärkning : En följd är både växande och avtagande $\Leftrightarrow$ den är konstant.

Sats 1 (Monoton Konvergens sats). *En monoton, begränsad följd av reella tal konvergerar.*

BEVIS : Gavs i föreläsningen. Se också GLO, s. 13.

Sats 2 (Bolzano-Weierstrass egenskap). *En godtycklig följd av reella tal har en oändlig monoton delföljd.*

BEVIS : Gavs i föreläsningen. Se också GLO, s. 13.

De här två satserna tillsammans ger

Sats 3 (Bolzano-Weierstrass sats). *Varje begränsad följd av reella tal har en konvergent delföljd.*

BEVIS : Följer omedelbart från Satser 1 och 2. Se också GLO, s. 14.

Med hjälp av Sats 3, tillsammans med följande lemma -

> LEMMA : Varje Cauchy följd av reella tal är begränsad.
> *Bevis.* Ska ges nästa gång.

- och lite mer arbete (som gjordes i föreläsningen) kan vi äntligen bevisa

Sats 4 (Cauchy konvergensprincip). *Varje Cauchy följd av reella tal konvergerar.*

OBS! Satser 3 och 4 kan enkelt generaliseras till följder i $\mathbf{R}^n$ (som görs i GLO, s. 14-15). Den ska vi prata om också nästa gång.

10 Sammandrag av föreläsning 10 Tisdag 5/9/99

Definition 1 (s. 17 har en annorlunda, och inte så vanlig, definition) : Låt $D \subseteq \mathbf{R}^n$ och $a \in D$. D kallas för en *omgivning* till a om $B(a, r) \subseteq D$ för någon $r > 0$.

Definition 2 (s. 18) : $D \subseteq \mathbf{R}^n$ kallas för en *öppen* mängd om den är en omgivning till var och en av dess punkter.

Definition 3 (*relativ topologi*) (finns inte i boken) : Låt $X \subseteq \mathbf{R}^n$ och $Y \subseteq X$. Y kallas för *öppen i* X om $Y = X \cap U$ för någon öppen delmängd U till $\mathbf{R}^n$.

'Unless otherwise stated', alla mängder betraktas som delmängder till någon $\mathbf{R}^n$. När det blir en situation där det passar bättre att betrakta våra mängder som delmängder till någonting annat, och att använda den relativa topologin, då ska jag berätta för er.

Proposition 1. *(i) ϕ är öppen.*
(ii) $\mathbf{R}^n$ själv är öppen.
(iii) En godtycklig union av öppna mängder är öppen.
(iv) En snitt av ändliga många öppna mängder är öppen.

Bevis : Gavs idag. Se också GLO, s. 21.

Anmärkning (inte i boken) : Låt X vara en godtycklig mängd och $\mathbf{T}$ en samling av delmängder till X. $\mathbf{T}$ kallas för en *topologi* på X om
(i) $\phi \in \mathbf{T}$.
(ii) $X \in \mathbf{T}$.
(iii) En godtycklig union av mängder i $\mathbf{T}$ är också i $\mathbf{T}$.
(iv) En snitt av ändliga många mängder i $\mathbf{T}$ är också i $\mathbf{T}$.
Paret $(X, \mathbf{T})$ kallas för ett *topologiskt rum*.

Definition 4 (s. 17) : Låt $D \subseteq \mathbf{R}^n$. En punkt $a \in D$ kallas för en *inre punkt* i D om D är en omgivning till a.

En punkt $a \in \mathbf{R}^n$ kallas för en *rand punkt* till D om varje omgivning till

a innehåller punkter från både D och D^c.

En punkt $a \in D^c$ kallas för en *yttre punkt* till D om den är en inre punkt till D^c.

NOTATION : Mängden av alla inre punkter i D kallas för *inren* till D och betecknas $\mathrm{Int}(D)$.

Mängden av alla rand punkter till D kallas för *randen* till D och betecknas ∂D.

Proposition 2. *För varje delmängd $D \subseteq \mathbf{R}^n$ får vi att*

(i) Int(D) är en öppen mängd i $\mathbf{R}^n$. Den är den storsta delmängden till D som är öppen i $\mathbf{R}^n$.

(ii) $\mathbf{R}^n =$ Int$(D) \sqcup \partial D \sqcup$ Int(D^c), där $\sqcup$ betecknar en disjunkt union.

BEVIS : Gavs idag. Finns inte i boken i en enda plats, men ... (blah!) ...

DEFINITION 5 (s. 18 har en annan definition, som svarar mot villkor (iii) i nästa proposition) : En delmängd D till $\mathbf{R}^n$ kallas för *sluten* om gränsvärdet till varje Cauchy följd av punkter i D ligger också i D.

Proposition 3. *Följande är ekvivalenta :*

(i) D är sluten.

(ii) D^c är öppen.

(iii) $D \supseteq \partial D$.

BEVIS : Gavs idag. Se också GLO, s. 18-19.

Nu ska vi ge två olika definitioner av 'kompakthet', som är det viktigaste begreppet i den här delen av kursen. Vår första definition låter abstrakt, men kan generaliseras till allmäna topologiska rum. Vår andra definition är mycket mer intuitivt. Den så kallade *Heine-Borel satsen* (se nere) säger att de två definitionerna är ekvivalenta.

DEFINITION 6 (s. 26) : Låt $X \subseteq \mathbf{R}^n$ och $\{Y_\lambda\}_{\lambda \in \Lambda}$ en familj av öppna mängder i X. Om $X = \bigcup_{\lambda \in \Lambda} Y_\lambda$ säger vi att de Y_λ *täcker* X, eller att de gör en *öppen täcke* av X.

Om $\Theta \subseteq \Lambda$ så att $X = \bigcup_{\lambda \in \Theta} Y_\lambda$, säger vi att de Y_λ, för $\lambda \in \Theta$, gör en *öppen deltäcke* till täcken som består av alla Y_λ $(\lambda \in \Lambda)$.

DEFINITION 7 (inte i boken) : En delmängd $D \subseteq \mathbf{R}^n$ kallas för *kompakt* om varje öppen täcke till D har en ändlig deltäcke.

DEFINITION 8 (s. 3) : En delmängd $D \subseteq \mathbf{R}^n$ kallas för *begränsad* om det existerar $K > 0$ så att $|x| \leq K$ för alla $x \in D$.

DEFINITION 9 (s. 23) : En delmängd $D \subseteq \mathbf{R}^n$ kallas för *kompakt* om D är sluten och begränsad.

DEFINITION 10 (denna terminologi används inte i boken) : En delmängd $D \subseteq \mathbf{R}^n$ kallas för *följdlig kompakt (eng.: sequentially compact)* om varje följd i D har en delföljd som konvergerar till en punkt i D.

SATS (Heine-Borel sats). *Följande är ekvivalenta, för en delmängd D till $\mathbf{R}^n$:*

(i) D är kompakt, med avseende på definition 7.

(ii) D är kompakt, med avseende på definition 9.

(iii) D är följdlig kompakt.

BEVIS : Nästa gång, ska vi bevisa att $(ii) \Rightarrow (iii)$. Se också GLO, s. 23. Det är lätt att se att $(i) \Leftrightarrow (ii)$. Det är mycket svårare att bevisa att $(ii) \Rightarrow (i)$ - se GLO, s. 27.

11 Sammandrag av föreläsning 11 Torsdag 7/9/99

Förra gången, hade vi kvar att bevisa

Sats (Heine-Borel). *Låt X vara en delmängd till någon $\mathbf{R}^n$. De följande är ekvivalenta :*

(i) varje öppen täcke till X har en ändlig deltäcke.

(ii) X är sluten och begränsad.

(iii) varje följd i X har en delföljd som konvergerar till en punkt i X.

BEVIS : Vi bevisade bara att (ii) $\Leftrightarrow$ (iii). Idéen för beviset är densamma som i bokens bevis, s. 23.

Kontinuerliga funktioner i Euklidiska rum

Hittills, har vi antagit att våra kontinuerliga funktioner hade öppna definitionsmängder. Men i många situationer, vi vill betrakta kontinuerliga funktioner med oöppna definitonsmängder, t.ex. en sluten intervall i $\mathbf{R}$ eller, mer allmänt, en kompakt mängd i någon $\mathbf{R}^n$. För att definiera kontinuitet för sådana funktioner, använder vi den relativa topologin (def. 1, nere). Satsen därefter ger också en ren 'topologisk' definition (del (iii) av satsen) av kontinuitet.

DEFINITION 1 : Låt $f : X \subseteq \mathbf{R}^n \to \mathbf{R}^m$. f sägs vara *kontinuerlig* vid en punkt $a \in X$ om, för varje $\epsilon > 0$, det finns $\delta > 0$ så att

$$x \in B(a, \delta) \cap X \Rightarrow f(x) \in B(f(a), \epsilon). \tag{95}$$

Ytterligare, f sägs vara *kontinuerlig på X* om den är kontinuerlig vid varje punkt i X.

Tillämpningar av nästa sats (särskilt del (iii)), som ger flera karakteriseringar av kontinuitet, finns ofta på de gamla tentorna. Först, lite notation för avbildningar mellan mängder :

NOTATION : Låt X, Y vara (godtyckliga) mängder och $f : X \to Y$ en (godtycklig) funktion.

(i) För $A \subseteq X$, definierar vi $f(A) :\stackrel{\text{def}}{=} \{f(x) : x \in A\}$.
(ii) För $B \subseteq Y$, definierar vi $f^{-1}(B) :\stackrel{\text{def}}{=} \{x \in X : f(x) \in B\}$.

Notera följande :

(a) $f^{-1}(f(A)) \supseteq A$ med likhet om f är injektiv.
(b) $f(f^{-1}(B)) \subseteq B$ med likhet om f är surjektiv.
(c) $f(\cup U_\lambda) = \cup f(U_\lambda)$, för en familj av delmängder $U_\lambda \subseteq X$.
(d) $f^{-1}(\cup V_\lambda) = \cup f^{-1}(V_\lambda)$, för en familj av delmängder $V_\lambda \subseteq Y$.
(e) Också $f^{-1}(\cap V_\lambda) = \cap f^{-1}(V_\lambda)$.
(f) $f^{-1}(B^c) = [f^{-1}(B)]^c$ för varje $B \subseteq Y$.

Nu kan vi ställa fram

Sats. *Låt $X \subseteq \mathbf{R}^n$, $Y \subseteq \mathbf{R}^m$, och $f : X \to Y$. Då är följande ekvivalenta :*
(i) f kontinuerlig på X.
(ii) Om (x_k) är en följd i X som konvergerar till en punkt $a \in X$, då konvergerar följden $(f(x_k))$ till $f(a) \in Y$.
(iii) $f^{-1}(V)$ är öppen i X för varje öppen delmängd $V \subseteq Y$.
(iv) $f^{-1}(V)$ är sluten i X för varje sluten delmängd $V \subseteq Y$.
(v) $f(\bar{A}) \subseteq \overline{f(A)}$, för varje $A \subseteq X$.

Bevis : Vi bevisade ekvivalensen för bara (i) - (iv), och struntade i (v), som ges inte i boken. Vårt bevis liknar det på s. 36-7 i boken.

Vårt mål för resten av dagens, och en del av nästa, föreläsning är att bevisa generaliseringar av två välkända satser från envariabelanalys (obs! dessa satser bevisas inte i envariabelanalysens kurs, och faktum är att deras bevis är inte som helst lättare än de mer allmäna versionerna). Alltså, kom ihåg följande resultater :

Sats A (Weierstrass' lemma). *En kontinuerlig reellvärd funktion definierad på en sluten intervall i $\mathbf{R}$ antar ett största och minsta värde.*

Sats B (Mellanliggandevärde sats). *Låt $f : X \subseteq \mathbf{R} \to \mathbf{R}$. Om $a, b \in f(X)$ med $a < b$, då är hela intervallen $[a, b] \subseteq f(X)$.*

Den lämpliga generaliseringen av Sats A är

Sats A'. *Låt $f : X \subseteq \mathbf{R}^n \to \mathbf{R}^m$, där X är kompakt och f kontinuerlig. Då är $f(X)$ kompakt.*

BEVIS : Vi gav ett bevis som utnyttar den abstrakta definitionen av kompakthet (varje öppen täcke har en ändlig deltäcke). Fördelen med detta bevis är att det är 'topologiskt' och kan generaliseras. I boken (s. 41) finns ett bevis, som ocksåär helt o.k., som använder den intuitiva definitionen av kompakthet (sluten och begränsad).

Vi visade också hur Weierstrass' lemma följer direkt från Sats A' (+ supremum/infimum axiomet). Se GLO, s. 41.

Att bevisa Sats B och dess generalisering behöver ganska mer 'set-up' arbete.

DEFINITION : En delmängd $X \subseteq \mathbf{R}^n$ kallas för *sammanhängande* om de enda delmängder till X som är både öppna och slutna **i X** är ϕ och X själv.

Ekvivalent, X är *sammanhängande* om den kan inte skrivas som en union av två disjunkta, icke-toma, öppna delmängder (där öppen betyder öppen i X).

OBS! Denna definition av sammanhängdehet är inte lika med bokens definition. Den senare är, faktist, en definition av begreppet *bågvis sammanhängdehet*. Varje båg. samm. mängd är samm. (kan lätt bevisas), men det finns samm. mängder som inte är båg. samm. (t.ex. Simmons, 'Introduction to topology and modern analysis', kap. 33). Dock, sådana mängder är ganska konstiga, och i de flesta vanliga situationerna, de två begreppen är ekvivalenta. Men jag tror det är bättre att berätta sänningen för er i alla fall !

Del (ii) av följande sats är det svåra steget (av två) i beviset av Sats B och dess generalisering.

Sats. (a) *Varje $\mathbf{R}^n$ är sammanhängande.*

(b) *En delmängd till $\mathbf{R}$ är sammanhängande om och endast om den är en intervall, dvs en av*

$$[a,b],(a,b),[a,b),(a,b],(-\infty,c],(-\infty,c),[c,+\infty),(c,+\infty). \qquad (96)$$

BEVIS : Vi gav ett komplett bevis av (ii). Det exakta samma beviset fungerar för bågvis sammanhängande delmängder till $\mathbf{R}$ och finns i boken, s. 46 (notera att det här betyder att det finns inga av dessa konstiga samm./ej båg. samm. mängder i $\mathbf{R}$: de finns ju i $\mathbf{R}^2$).

Samma idé används i beviset av (i). Jag skrev inte ner alla detaljer, och de behövs inte.

Vi ska ta steg 2 i beviset av Sats B (se GLO, s. 44) nästa gång.

12 Sammandrag av föreläsning 12 Tisdag 12/10/99

Notera att den mellanliggandevärde satsen kan reformuleras som

'*Om* $f : \mathbf{R} \to \mathbf{R}$ *är kontinuerlig och* $I \subseteq \mathbf{R}$ *är en intervall, då är* $f(I)$ *också en intervall*'.

Sista gång, bevisade vi en sats som sa att en delmängd till $\mathbf{R}$ är sammanhängande om och endast om den är en intervall. Alltså, följande sats generaliserar den m.l.v. satsen :

Sats. *Låt* $f : X \subseteq \mathbf{R}^n \to \mathbf{R}^m$ *vara kontinuerlig. Om* X *är sammanhängande, så också är* $f(X)$.

Denna sats säger att sammanhängandehet (liksom kompakthet) är en topologisk egenskap.

Bevis : Gavs i föreläsningen. Finns inte i GLO.

Ett begrepp som är närbesläktad med sammanhängdehet är *bågvis sammanhängdehet*. Detta har vi redan definierat i föreläsning 4. Den är ingen överraskning att vi har

Proposition. *Varje bågivs sammanhängande delmängd till någon* $\mathbf{R}^n$ *är sammanhängande.*

Bevis : Vi skisserade beviset (finns inte i boken). Idéen är att om X är b.s.h. men inte s.h., får vi en avkoppling (disconnection) av den slutna intervallen $[0, 1]$ - motsägelse !

Den omvänt propositionen stämmer i $\mathbf{R}$ (ty en s.h. mängd är en intervall, som är triviallt b.s.h.), men inte i högre dimensioner. För ett exempel av en s.h. delmängd till $\mathbf{R}^2$ som inte är b.s.h., se t. ex. Simmons, 'Introduction to Topology and Modern Analysis', kap. 33. Sådana mängder är ganska konstiga.

Bågvis sammanhängdehet är också en topologisk egenskap, dvs

Sats. *Om $f : X \subseteq \mathbf{R}^n \to \mathbf{R}^m$ är kontinuerlig, och X bågvis sammanhängande, då $f(X)$ är också bågvis sammanhängande.*

Bevis : Disktuerades inte. Se GLO, s. 44.

Topologiska egenskaper och avbildningar

Låt D, S vara delmängder till någon $\mathbf{R}^n$ och $\mathbf{R}^m$ resp. En naturlig och viktig fråga i topologi är

'Finns det en kontinuerlig surjektion $f : D \to S$?'

Man kan inte förvänta sig att det blir en allmän metod för att lösa detta problem. Men ett nödvändigt villkor för existens av en sådan avbildning är att S äger (possesses) varje topologisk egenskap som gör D. Alltså, om D är kompakt/sammanhängande/bågvis sammanhängande, så också är S. Det finns ingen komplett lista av topologiska egenskaper, men om man tar några av de vanligaste, och ser att S äger var och en av dem som gör D, man börjar tro att en kontinuerlig surjektion existerar. I detta fall, det återstår att beskriva en explicit f. Detta brukar innehålla lite geometrisk/topologisk resonnemang. Oftast, man bygger upp f i flera steg, som är tillåtat eftersom en sammansattning av kontinuerliga funktioner är kontinuerlig. För att beskriva hela f, det räcker att beskriva varje steg. Ofta, en explicit formul till f är mycket komplicerad att skriva ner, men det behövs inte om varje steg har en klar tolkning.

Exempel : Vi löste några problem från gamla tentor, dvs 980318 (nr. 4) och 980818 (nr. 4). Kompletta lösningar till dessa och sex andra tentor ska delas ut på torsdag.

13 Sammandrag av föreläsning 13 Torsdag 14/9/99

Idag, ska vi diskutera *optimerings problem*, dvs matematiska problem där vi försöker hitta de största och minsta värden, om de existerar, antagit av en reellvärd funktion $f : D \subseteq \mathbf{R}^n \to \mathbf{R}$. Detta är ett globalt problem, dvs vi söker globala extrema till f, som oftast har ingen lösning. Emellertid, finns det minst tre allmäna situationer där existens av en lösning är garanterat och vi har en 'algoritmisk' procedur för att ange den. Alla tre situationer har viktiga praktiska tillämpningar. I två av dem, inskränkar vi definitionsmängden D. I den tredje, gör vi några förutsättningar om f's asymptotisk uppförande.

Fall 1 : D är kompakt

Vi har bevisat en sats som säger att varje kontinuerlig reellvärd funktion med en kompakt definitionsmängd antar ett max. och min. värde. Om f är dessutom (minst) C^1, får vi följande 'algoritm' för att hitta dessa extremvärde :

Proposition 1. *Låt $f : D \subseteq \mathbf{R}^n \to R$ vara en C^1 funktion där D är kompakt. Antag att $f = f(x_1, ..., x_n)$ antar en global extremum (max. eller min.) i punkten P. Då, antingen*

(i) $P \in Int(D)$ och P är en kritisk punkt till f, dvs

$$\frac{\partial f}{\partial x_i}(P) = 0, \quad i = 1, ..., n. \tag{97}$$

eller

(ii) $P \in \partial D$.

Detta resultat kan rättvist beskrivas som en (induktiv) algoritm för att hitta f's extremvärde eftersom ∂D är en mängd av dimension högst $n-1$, som kan paramateriseras med högst $n-1$ C^1-parameter, så att $f|_{\partial D}$ kan betraktas som en $(n-1)$-variabel funktion vars extremvärde undersökes[1].

[1] Denna är en ren intuitiv förklaring. Vi skall inte försöka beskriva de exakta villkoren som garanterar att ∂D kan beskrivas med högst $n-1$ C^1-parameter.

Vi diskuterade följande två exempel. Exempel 2 kan reduceras till exempel 1 :

EXEMPEL 1 : Maximisera $f(x,y) = xy(108 - 2x - 2y)$ i det triangulära området begränsad av (bounded by) den x-axeln, y-axeln och linjen $x + y = 54$.

EXEMPEL 2 : Man vill skicka ett rektangulär paket i posten. Summan av paketets längd och omkrets (perimeter/girth) måste vara ≤ 108cm. Vad är paketets max. möglig volym ?

Följande exempel ska diskuteras i tisdags lektion. Den nya idéen här är att byta till polära koordinator för att undersöka f på ∂D.

EXEMPEL 3 : Ange max/min till $f(x,y) = x^2 + y^2 - 2x - 4y$ i den slutna skivan av radie 3 kring origon.

Fall 2 : $\mathbf{f} \to \mathbf{0}$ när $|\mathbf{x}| \to \infty$

Vi bevisade följande (lätt) observation :

Proposition 2. *Låt $f : \mathbf{R}^n \to \mathbf{R}$ vara C^1 och antag att $f(\vec{x}) \to 0$ när $|\vec{x}| \to \infty$. Då,*

(i) om det finns minst en punkt $\vec{a}$ så att $f(\vec{a}) > 0$, då har f en global maximum. Om denna antas i punkten P, då är P en kritisk punkt till f.

(ii) om det finns minst en punkt $\vec{a}$ där $f(\vec{a}) < 0$, då har f en global minimum. Om denna antas i punkten Q, då är Q en kritisk punkt till f.

Denna proposition har följande 'dual' (svenska ?)

Proposition 3. *Låt $f : \mathbf{R}^n \to \mathbf{R}$ vara C^1.*

(i) Om $f(\vec{x}) \to +\infty$ när $|\vec{x}| \to \infty$, då har f en global minimum. Om denna antas i punkten P, då är P en kritisk punkt till f.

(ii) Om $f(\vec{x}) \to -\infty$ när $|\vec{x}| \to \infty$, då har f en global maximum. Om

denna antas i punkten Q, då är Q en kritisk punkt till f.

Vi diskuterade följande exempel :

EXEMPEL 4 : Bestäm de största och minsta värden till $f(x,y) = (2x^2 + y^2)\exp(-x^2 - y^2)$.

EXEMPEL 5 : Vilken punkt i planet $x + 2y - z = 3$ är närmast origon ?

OBS! (i) Jag tror att jag gjorde ett räknefel i lösningen till exempel 5, så vi ska göra om den i lektionen på tisdag.
(ii) Exempel 5 kan också lösas med de så kallade *Lagrange multiplikatorerna*, som ska inleds nästa gång.

14 Sammandrag av föreläsning 14 Tisdag 19/9/99

Idag, förtsätter vi med optimerings problem. Vi har ett fall kvar att diskutera.

Fall 3 : Bivillkor

Som vanligt, börjar vi med en C^1-funktion $f : D \subseteq \mathbf{R}^n \to \mathbf{R}$. Vi gör inga förutsättningar om definitionsmängden D, som är ofta hela $\mathbf{R}^n$. Men, i föreliggande fall, är vi inte interesserad av globala extrema till f, men bara av extrema i skärningen mellan D och en eller fler hyper-ytor (dvs, $(n-1)$-dimensionella delmängder) i $\mathbf{R}^n$. Vi antar att varje sådan yta kan beskrivas

$$g(x_1, ..., x_n) = 0 \tag{98}$$

för någon n-variabel C^1-funktion g. I allmänhet, skärningen mellan k sådana ytor ska vara en $(n-k)$-dimensionell mängd. Alltså, i Fall 3, söker vi de extremvärden till en n-variabel C^1-funktion f i en $(n-k)$-dimensionell delmängd till dess definitionsmängd D som beskrivas som skärningen mellan D och k stycke $(n-1)$-dimensionella C^1-hyperytor i $\mathbf{R}^n$[2].

Vi gav följande sats (PB, 4.3) utan bevis (som inte behövs) :

Sats 1. *Låt $f, g : \mathbf{R}^2 \to \mathbf{R}$ vara C^1-funktioner. Antag att ett extremvärde till f längs kurvan $g(x,y) = 0$ antas i punkten (a,b), och att (a,b) ligger i inren till f's definitionsmängd*[3]*. Då är vektorerna $\nabla f(a,b)$ och $\nabla g(a,b)$ parallella.*

Denna slutsats är oanvändbar endast om $\nabla g(a,b) = \vec{0}$, ty varje vektor är parallell med den nolla vektoren. Annars, satsen säger att det finns en konstant $\lambda \in \mathbf{R}$ så att

$$\nabla f(a,b) = \lambda \nabla g(a,b). \tag{99}$$

[2]Den är en interessant, men ganska abstrakt fråga i *differential geometry* vilka $(n-k)$-dimensionella delmängder till $\mathbf{R}^n$ kan beskrivas som en skärning av k stycke (C^1, C^∞ mm)-hyperytor. En svårighet med att ge en precis formulering av detta problem är att ge en precis formulering av termen 'dimension', för en godtycklig delmängd till ett Euklidisk rum.

[3]Detta är ett tekniskt villkor som oftast kan ignoreras.

Konstanten λ kallas för en *Lagrange multiplikator*. Eftersom två vektorer är lika omm alla motsvarande komponenter är lika, (2) ger de två ekvationerna

$$\begin{aligned} \frac{\partial f}{\partial x}(a,b) &= \lambda \frac{\partial g}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b) &= \lambda \frac{\partial g}{\partial y}(a,b). \end{aligned} \tag{100}$$

Dessa ekvationer, tillsammans med

$$g(a,b) = 0, \tag{101}$$

ger ett system av tre ekvationer för de tre obekanta λ, a, b som man hoppas kunna lösa (i det värsta fallet, man kan approximera lösningarna med hjälp av en dator).

OBS! Det händer ofta att systemet (100) och (101) har ingen lösning, så att optimerings problemet har ingen lösning heller. Detta är ingen överrakning eftersom, t.ex., kurvan $g(x,y) = 0$ brukar inte vara en kompakt mängd så att vi kan inte förvänta oss att funktionen f antar antingen ett störst eller minst värde längs den. I tillämpningar, å andra sidan, man brukar betrakta situationer där man vet a priori att ett extremvärde existerar. Alltså, ovanstående formuler tillåter man att hitta extremvärdet.

Sats 1 har följande generalisering till funktioner av n variabler, med k $(1 \leq k < n)$ bivillkor :

Sats 2. *Låt $f, g_1, ..., g_k : \mathbf{R}^n \to \mathbf{R}$ vara C^1-funktioner. Antag att ett extremvärde till f på den $(n-k)$-dimensionella skärningen mellan de hyperytorna $g_i(x_1, ..., x_n) = 0$ $(1 \leq i \leq k)$ antas i punkten $\vec{a} = (a_1, ..., a_n)$ och att $\vec{a}$ ligger i inren till f's definitionsmängd. Då är de $k+1$ vektorerna $\nabla f(\vec{a}), \nabla g_1(\vec{a}), ..., \nabla g_k(\vec{a})$ linjärt beroende, dvs det finns konstanter $\lambda_0, \lambda_1, ...\lambda_k \in \mathbf{R}$, inte alla noll (not all zero), så att*

$$\lambda_0 \nabla f(\vec{a}) + \sum_{i=1}^{k} \lambda_i \nabla g_i(\vec{a}) = \vec{0}. \tag{102}$$

Denna sats är oanvändbar om de k vektorerna $\nabla g_1(\vec{a}), ..., \nabla g_k(\vec{a})$ är redan linjärt beroende, så att vi kan välja $\lambda_0 = 0$ i (102). Annars, kan (102) skrivas som

$$\nabla f(\vec{a}) = \sum_{i=1}^{k} \mu_i \nabla g_i(\vec{a}), \tag{103}$$

för några k-tupel $(\mu_1, ..., \mu_k)$ av reella tal. Denna ekvation mellan vektorer i $\mathbf{R}^n$ ger följande n skalära ekvationer

$$\frac{\partial f}{\partial x_j} = \sum_{i=1}^{k} \mu_i \frac{\partial g_i}{\partial x_j}, \quad j = 1, ..., n. \tag{104}$$

Dessa, tillsammans med de k bivillkoren

$$g_i(\vec{a}) = 0, \quad i = 1, ..., k \tag{105}$$

ger ett system av $n+k$ ekvationer för de $n+k$ obekanta $a_1, ..., a_n, \mu_1, ..., \mu_k$.

OBS! De övanstående μ_i kallas också för Lagrange multiplikatorer. I det optimerings problemet, man är interesserad av att hitta punkten $\vec{a} = (a_1, ..., a_n)$ där extremvärdet antas. Man är inte interesserad av de exakta värden av de Lagrange multiplikatorerna. Därför, när man försöker lösa ett ekvation system av övanstående typ, är det bäst att eliminera de μ_i först, om mögligt.

Vi diskuterade följande exempel

Exempel 1 : Repitition av exempel 5, föreläsning 13, men denna gång använder vi Lagrange multiplikatorer.

Exempel 2 : Maximisera och minimisera $f(x, y) = x^2 y$ på enhetscirkeln.

OBS! Både dessa exempel visar att man kan slippa använda Lag. mult. om man kan hitta en explicit parameterisering av den $(n-k)$-dimensionella mängden där optimeringsproblemet ska lösas. T.ex., i exempel 2 kan vi parameterisera cirkeln med polära koordinator. I de flesta situationerna, emellertid, är det inte så lätt att hitta en sådan explicit parameterisering, och mycket mer effektivt att använda den allmäna multiplikatorer metoden.

Exempel 3 (AMGM olikheten): Maximisera $f(x_1, ..., x_n) = \prod_{i=1}^{n} x_i$, om varje $x_i \geq 0$ och $\sum_{i=1}^{n} x_i = 1$.

Exempel 4 (mer än ett bivillkor) : Vilken punkt på linjen som beskriver skärningen mellan de två planen $x + 2y + 3z = 6$ och $x + 3y + 9z = 9$ är närmast origon ?

(Lösning : $(\frac{81}{59}, \frac{121}{59}, \frac{9}{59})$).

OBS! I exempel 4, får man ett system av fem linjära ekvationer i fem obekanta, som har en unik lösning (a priori, lösningen är unik, ty det finns en unik punkt längs en rät linje närmast en punkt utanför linjen). Det är, faktiskt, lättare att glömma multiplikatorer och att lösa problemet genom att hitta en explicit parameterisering av den räta linjen. Alltså, kom ihåg att den multiplikatorer metoden ger en allmän metod för att lösa optimerings problem. Det följer inte att den ger den bästa metoden i varje situation.

Exempel 4 visar också att när man användar den multiplikatorer metoden, man brukar har kvar ett ganska stort system av ekvationer (normalt, icke-linjära) att lösa. I tillämpningar, man måste ofta nöja sig med en approximat lösning, och behöver en kraftig dator för att få den.

15 Sammandrag av föreläsning 15 Torsdag 21/10/99

Deifferentiation under integral tecken

I många situationer, särskilt i Fourier (harmonisk) analys och lösande av diffekvationer, träffar man funktioner definierad som en definit integral av någon annan funktion. I det enklaste fallet, har vi en funktion

$$F(s) = \int_a^b f(s,t)dt, \tag{106}$$

av den enda variabeln s där $f(s,t)$ är kontinuerlig i t (alltså kan integreras) för varje s. Antag nu att $F(s)$ är differentierbar. Det är naturligt att misstänka att

$$F'(s) = \int_a^b \frac{\partial f}{\partial s}(s,t)dt. \tag{107}$$

Lyckligtvis är Gud god (? God is good !) och (107) stämmer under några svaga villkor. Vi har

Sats 1. *Låt $f(s,t)$ vara kontinuerlig och C^1 med avseende på s. Då är funktionen $F(s)$ i (1) differentierbar, för godtyckliga reella tal a, b och (107) stämmer.*

Bevis : Tekniskt - någonting som man läser och glömmer omedelbart. Därför struntar vi i det.

Denna sats har (minst) två generaliseringar. I den första, skapar vi en lite mer komplicerad funktion av s genom att välja godtyckliga differentierbara funktioner $a(s), b(s)$ som integralens ändpunkter i (106), i stället för konstanter a, b. I detta fall, kan man bevisa följande

Sats 2. *Låt $f(s,t)$ vara kontinuerlig och C^1 med avseende på s. Låt $a(s), b(s)$ vara differentierbara funktioner av s. Då är*

$$F(s) = \int_{a(s)}^{b(s)} f(s,t)dt \tag{108}$$

en differentierbar funktion av s och

$$F'(s) = \int_{a(s)}^{b(s)} \frac{\partial f}{\partial s}(s,t)dt + b'(s) \cdot f(s, b(s)) - a'(s) \cdot f(s, a(s)). \tag{109}$$

Bevis : Gavs inte, men följer från Sats 1 + Kalkyls Fondamental Sats + 1-variabel kedjeregeln. Notera att (109) reduceras till (107) om $a(s) = a$ och $b(s) = b$ (konstanter).

För den andra generaliseringen, betraktar vi en integral över en oändlig intervall, dvs en integral av formen

$$F(s) = \int_a^\infty f(s,t)dt, \tag{110}$$

där vi förutsätter att integralen konvergerar för alla s i någon öppen delmängd I till $\mathbf{R}$. Vi hoppas kunna fortfarande skriva

$$F'(s) = \int_a^\infty \frac{\partial f}{\partial s}(s,t)dt, \tag{111}$$

under några svaga villkor. Det visar sig (? it turns out that) funktionen f måste satisfiera ett starkare villkor än i Sats 1, men fortfarande ett villkor som oftast uppfyllas. Vi har

Sats 3. *Låt $f(s,t)$ vara kontinuerlig och C^1 med avseende på s. Antag att $\int_a^\infty f(s,t)dt$ existerar för alla s i en öppen delmängd I till $\mathbf{R}$ och att, för varje kompakt delmängd K till I, det finns en kontinuerlig funktion $g_K(t)$ så att $|\frac{\partial f}{\partial s}(s,t)| \leq g_K(t)$ för alla $s \in K, t \geq a$ och $\int_a^\infty g_K(t)dt < \infty$. Då är funktionen F(s) i (110) differentierbar och (111) stämmer.*

Bevis : Tekniskt igen - gavs inte. Den terminologin som används för det nya villkoret som beskrivs i denna sats är att $\frac{\partial f}{\partial s}$ är *likformigt integrerbar (eng.: uniformly integrable)*. De funktioner g_K kallas för *majorerande funktioner*. Det är värd att notera att naturvetenskapliga förskare ofta använder ekvation (111) utan att kontrollera att de villkoren i sats 3 uppfyllas. Detta anledar ibland till ett falst resultat.

Exempel : Vi diskuterade exempel 1, s. 163 i boken och löste övning 5.1 i övningsboken. På tisdag, ska vi lösa 5.4 och 5.18 och 990816 f.7, 980318 f.7. Punkten med de flesta av dessa exempel är att uppvisa en mycket vacker tillämpning av differentiation under integral tecken, till lösande av många 1-variabel definita integraler där man kan inte skriva ner en explicit anti-derivata till integranden. En annan tillämpning av denna metod är till de så kallade *Fourier* och *Laplace transformationerna* av kontinuerliga funktioner, som används ofta i undersökningen av diffekvationer och i harmonisk (Fourier) analys. Vi hinner inte ge en bra inledning till dessa ämne ; diskussionen i boken är också långt ifrån komplett. Alltså, som de säger i New York, 'forget about it !'.

16 Sammandrag av föreläsning 16 Tisdag 26/10/99

En generalisering till av sats 1 från torsdag :

Sats 4. *Låt $f(s,t)$ vara en C^k-funktion av s och t, och C^{k+1} i variabeln s. Sätt*

$$F(s) := \int_a^b f(s,t)\, dt. \tag{112}$$

Då är $F(s)$ differentierbar minst $k+1$ gånger och

$$F^{(i)}(s) = \int_a^b \frac{\partial^i f}{\partial s^i}(s,t)\, dt, \tag{113}$$

för $1 \leq i \leq k+1$.