KLASSIFICERING AV STATIONÄRA PUNKTER - HESSIANEN

Detta dokument är ett supplement till matieralet i avsnitt 2.6 i PB och består av två avsnitt:

- Binära kvadratiska former
- Matrisformulering
- Generalisering till fler variabler

Materialet i det första avsnittet ska betraktas som examinerbart. Materialet i det andra och tredje avsnittet är "frivilligt", dock kan användas fritt vid undersökning av funktioner av fler än två variabler. Dessutom är det bra att vara bekant med ordet *Hessian(matris)*, som förklaras mot slutet.

I bokens så tacklas varje uppgift genom att göra en explicit kvadratkomplettering av uppgiftens kvadratiska form, men när antalet variabler överstiger två så saknas en uppenbar metod för att gå tillväga med detta. I synnerhet *Sylvesters sats*, som anges mot slutet på detta dokument, är ett sätt att ta sig runt detta problem.

Binära kvadratiska former. Se boken s. 99-100 för definitioner av termerna *lokalt maximum, lokalt minimum, lokal extrempunkt, stationär punkt.*

DEFINITION. En funktion $Q : \mathbb{R}^2 \to \mathbb{R}$ på formen

$$Q(x,\, y) = Ax^2 + 2Bxy + Cy^2, \quad A,\, B,\, C \in \mathbb{R}, \tag{1.1}$$

kallas för en *binär kvadratisk form.* Se boken s.102 för definitionerna av begreppen *positiv definit, negativ definit, indefinit, semidefinit* för binära kvadratiska former.

SATS 1. *För formen Q i (1.1) gäller:*
(i) Om $AC - B^2 > 0$ och $A > 0$ så är Q positiv definit.
(ii) Om $AC - B^2 > 0$ och $A < 0$ så är Q negativ definit.
(iii) Om $AC - B^2 < 0$ så är Q indefinit.
(iv) Om $AC - B^2 = 0$ så är Q semidefinit.

BEVIS. Om $A \neq 0$ så kan (1.1) skrivas om till

$$Q(x,\, y) = A\left[\left(x + \frac{B}{A}y\right)^2 + (AC - B^2)\left(\frac{y}{A}\right)^2\right]. \tag{1.2}$$

Om $AC - B^2 > 0$ så är uttrycket inom $[\ldots]$ alltid positiv (pga kvadraterna) och lika med noll endast då $x + \frac{B}{A}y = \frac{y}{A} = 0$, dvs endast då $(x, y) = (0, 0)$. Tecknet hos hela HL av (1.2) bestäms således av tecknet hos A. Detta bevisar (i) och (ii).

Om däremot $AC - B^2 < 0$ så kommer uttrycket inom $[\ldots]$ att kunna anta både positiva och negativa värden, som bevisar (iii) i fall $A \neq 0$. Man kan föra ett motsvarande resonemang om $C \neq 0$, bara låt x och y byta roller. Om $A = C = 0$ så är $AC - B^2 = -B^2$. Detta är aldrig positivt, och är lika med noll endast då $B = 0$. I så fall har vi nu $A = B = C = 0$ och Q är tydligen semidefinit, i enlighet med (iv). Om $A = C = 0$ och $B \neq 0$ så är $Q(x,y) = 2Bxy$, vilket kan uppenbarligen anta både positiva och negativa värden beroende på tecknen hos x och y. Då är (iii) helt bevisat.

Det som återsår är att betrakta fallet då $AC - B^2 = 0$, men minst ett av A och C är skilt från noll. Från symmetrin kan vi anta att $A \neq 0$ så att (1.2) gäller och blir till $Q(x, y) = A\left(x + \frac{B}{A}y\right)^2$. Det är klart att $Q(x, y)$ har alltid samma tecken som A, men $Q(x, y) = 0$ längs linjen $x + \frac{B}{A}y = 0$. Q är således semidefinit, i enlighet med (iv). ■

Anmärkning. De fyra fallen ovan täcker alla möjligheter för koefficienterna A, B, C ty om $A = 0$ så kan $AC - B^2$ inte vara strängt positiv.

Sats 2. *Låt $f : \mathbb{R}^2 \to \mathbb{R}$ och antag att $\in C^3$ för enkelhets skull. Låt $(a,\, b)$ vara en stationär punkt till f. Sätt $A = f_{xx}(a, b)$, $B = f_{xy}(a, b)$ och $C = f_{yy}(a, b)$. Då gäller:*
(i) Om $AC - B^2 > 0$ och $A > 0$ så är $(a,\, b)$ ett lokalt minimum för f.
(ii) Om $AC - B^2 > 0$ och $A < 0$ så är (a, b) ett lokalt maximum för f.
(iii) Om $AC - B^2 < 0$ så är (a, b) en sadelpunkt för f.

Anmärkning. Om $AC - B^2 = 0$ så ger Sats 2 ingen information om den stationära punktens karaktär. I sådana fall måste man antingen titta längre ut i Taylorutvecklingen av f i $(a,\, b)$ eller använda annan kunskap man har om just denna specifika f.

Bevis av Sats 2. Satsen följer direkt från Sats 2.6.12 i boken plus Sats 1 ovan.

Matrisformulering. Sätt $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, $M = \begin{bmatrix} A & B \\ B & C \end{bmatrix}$. Då kan (1.1) skrivas i matrisform som

$$Q(\boldsymbol{x}) = \boldsymbol{x}^T M \boldsymbol{x}. \tag{1.3}$$

Det kompelxa talet λ sägs vara ett *egenvärde* till matrisen M om $\det(M - \lambda I_2) = 0$. Skriver man upp detta explicit så blir det en kvadratisk ekvation för λ (som kallas för M:s *karakteristiska ekvation*):

$$\lambda^2 - (A + C)\lambda + (AC - B^2) = 0. \tag{1.4}$$

Man kan kontrollera att de två lösningarna λ_1 och λ_2 (som kan sammanfalla) är alltid reella tal och att det finns följande samband med Sats 1:

Sats 3. *(i) Om $\lambda_1 > 0$ och $\lambda_2 > 0$ så är Q positiv definit.*
(ii) Om $\lambda_1 < 0$ och $\lambda_2 < 0$ så är Q negativ definit.
(iii) Om $\lambda_1 > 0$, $\lambda_2 < 0$ eller vice versa så är Q indefinit.
(iv) Om $\lambda_1\lambda_2 = 0$ så är Q semidefinit.

Dessutom går det att bevisa (standard linjär algebra) att det finns en *ortogonalmatris* P (dvs en matris sådan att $P^T = P^{-1}$) sådan att $M = P^T D P$, $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. Detta kallas för en *ortogonaldiagonalisering* av matrisen M. Den medför i sin tur att, om vi utför variabelbytet $\begin{bmatrix} z \\ w \end{bmatrix} = \boldsymbol{y} := P\boldsymbol{x}$, så omvandlas (1.1) till

$$Q(z,\, w) = \lambda_1 z^2 + \lambda_2 w^2. \tag{1.5}$$

Konstatera att Sats 3 följer lätt från denna framställning av Q.

Generalisering till fler variabler. En kvadratisk form i n variabler är en funktion på formen

$$Q(x_1,\, \dots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{i<j} 2a_{ij} x_i x_j, \quad a_{ij} \in \mathbb{R}. \tag{1.6}$$

Sätter vi $\boldsymbol{x} = [x_1\, x_2\, \dots\, x_n]^T$ och låter M vara den symmetriska matrisen som har talen a_{ij} ovanför diagonalen så kan (1.6) skrivas i matrisform som i (1.3). Den karakteristiska ekvationen till M blir nu en n:te grads polynomekvation. Det finns en sats i linjär algebra som säger att alla egenvärdena till en symmetrisk matris är reella[1] och att man kan alltid hitta en ortogonal-diagonalisering av M och således ett variabelbyte sådant att Q kan framställas som

$$Q(y_1, \ldots, y_n) = \sum_{i=1}^{n} \lambda_i y_i^2. \tag{1.7}$$

Från (1.7) följer nu lätt:

SATS 4. *Låt Q vara en kvadratisk form i n variabler med tillhörande egenvärden $\lambda_1, \ldots, \lambda_n$. Då gäller:*
(i) Om alla $\lambda_i > 0$ så är Q positiv definit.
(ii) Om alla $\lambda_i < 0$ så är Q negativ definit.
(iii) Om $\lambda_i > 0$ för något i och $\lambda_j < 0$ för något j, så är Q indefinit.
(iv) Q är positiv semidefinit (resp. negativ semidefinit) om $\lambda_i \geq 0$ för varje i (resp. $\lambda_i \leq 0$ för varje i) och $\lambda_i = 0$ för minst ett i.

Sats 4 är svårt att tillämpa direkt för $n > 2$ ty man måste först lösa en n:te grads ekvation för att hitta egenvärdena. I praktik är detta sällan ett problem ty det är lätt att uppskatta egenvärdena med hjälp av en datorberäkning, och det är bara tecknen hos egenvärdena som spelar roll - förutom när ett av egenvärdena är noll, men detta är ekvivalent med att $\det(M) = 0$ och en determinant kan beräknas mer effektivt via Gausselimination. Vill man ha ett exakt kriterium som kan tillämpas direkt dock kan man använda följande sats, vars bevis vi avstår. OBS! dock att då $n = 2$ så är satsen ekvivalent med Sats 1 ovan.

SYLVESTERS SATS. *Låt Q vara en kvadratisk form i n variabler, med tillhörande symmetriska matris M. För $i = 1, \ldots, n$, låt M_i vara $i \times i$ matrisen i M:s övre vänstre hörn, och låt $d_i := det(M_i)$. Då gäller:*
(i) Om $d_i > 0$ för varje i, så är Q positiv definit.
(ii) Om $d_i > 0$ för jämna i medan att $d_i < 0$ för udda i, så är Q negativ definit.
(iii) *Om $d_n \neq 0$ men varken (i) eller (ii) gäller, så är Q indefinit.*
(iv) Om $d_n = 0$ så är Q varken positiv definit eller negativ definit. Q kan vara antingen semidefinit eller indefinit.

Vid tillämpning till klassificering av stationära punkter hos en funktion $f(x_1, \ldots, x_n)$ av n variabler så väljer man $a_{ij} = f_{x_i x_j}(\boldsymbol{a})$ i (1.6), där $\boldsymbol{a}$ är den stationära punkten ifråga. Tillhörande matrisen M kallas för *Hessianen* till f (i punkten $\boldsymbol{a}$). Speciellt viktigt som alltid är fallet $n = 2$: om $f = f(x, y)$ så är $\text{Hess}(f) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$.

ANMÄRKNING. Under diskussionen ovan har vi antagit att $f \in C^3$. Men man kan bevisa att teorin går igenom så länge $f \in C^2$. Vi avstår från detaljer.

[1]Denna sats har användning inom kvantmekanik, där egenvärdena tolkas som energinivåer.