Estudiem la funció $f(x_1, \dots, x_n)$ en un domini obert on existeixen, i són contínues, totes les seves derivades parcials fins l'ordre que sigui necessari.

· PUNT CRÍTIC (o ESTACIONARI): $a(a_1, \dots, a_n)$ és punt estacionari d' f quan totes les derivades parcials (primeres) d' f s'anul·len alhora en a.

· TAYLOR, vector GRADIENT i matriu HESSIANA: tindrem en ment tota l'estona la <u>fórmula de Taylor</u> de segon ordre al voltant d' a,

$$f(x) = f(a) \; + \sum_i \partial_i f \cdot u^i \; + \frac{1}{2}\sum_{i,j} \partial_{ij}^2 f \cdot u^i u^j \; + \cdots$$

sent-hi:

$$u^i = (x_i - a_i), \qquad \partial_i f = \frac{\partial f}{\partial x_i} = f_i, \qquad \partial_{ij}^2 f = \partial_i(\partial_j f) = \frac{\partial^2 f}{\partial x_i \partial x_j} = f_{ij}$$

i entenent que avaluem en a totes les derivades parcials involucrades. Per tant, ens apareixen les components del <u>vector gradient</u> d' f en el punt a:

$$\operatorname{grad} f = \nabla f = (\partial_1 f, \dots, \partial_n f)$$

i també les de la <u>matriu hessiana</u> d' f en el punt a:

$$H = \begin{pmatrix} \partial_{11}^2 f & \cdots & \partial_{1n}^2 f \\ \vdots & \ddots & \vdots \\ \partial_{n1}^2 f & \cdots & \partial_{nn}^2 f \end{pmatrix} = \begin{pmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nn} \end{pmatrix} \quad \Leftrightarrow \quad H_{ij} = f_{ij}$$

· CONDICIÓ NECESSÀRIA d'EXTREM:

Per a tenir un extrem (màxim o mínim) en a, anàlogament a funcions d'una variable, si ens situem en $x = a$ i imaginem que ens desplacem una mica, el valor de la funció no ha de canviar a primer ordre, $f(x) \cong f(a)$. Això només pot ser veritat, mirant la fórmula de Taylor, quan totes les components del gradient s'anul·lin alhora; és a dir, quan el punt sigui estacionari. En conclusió,

«per a que a sigui extrem, es condició necessària que sigui estacionari».

· CLASSIFICACIÓ de PUNTS ESTACIONARIS amb AUTOVALORS:

Una vegada sabem que a és estacionari, podria ser màxim, mínim o "punt sella". El significat d'aquests tres diferents tipus de punts s'entén un altre cop si ens situem en $x = a$ i imaginem que ens desplacem una mica: si el valor de la funció sempre decreix, $f(x) < f(a)$, llavors estarem en presència d'un màxim. Si sempre creix, a serà mínim. Si depenent de la direcció en què ens allunyem d' a la funció creix o decreix, llavors és un punt sella.

Tot seguit anem a estudiar el comportament de la funció analitzant a segon ordre en Taylor com varia quan ens desplacem des d' a. Trobarem condicions necessàries i suficients per a creixement (o decreixement) a ordre dos independentment de la direcció en què ens allunyem. Això voldrà dir que tindrem mínim (o màxim) en a, però no serà l'única manera possible de tenir mínim (o màxim) en a. Això és així perquè la funció podria, per exemple, no variar a primer i segon ordre i créixer a ordres superiors, com ocorre en una dimensió per a la funció $f(x) = x^4$ en l'origen (que és mínim: no varia a ordres primer, segon i tercer, però si arribem a quart sempre creix). És a dir: aquí anem a estudiar la caracterització d'un tipus d'extrems molt importants, que en direm "extrems a segon ordre", però que no són els únics possibles. És important assenyalar, per tant, que les condicions d'extrem que anem a trobar fent aquest estudi "a segon ordre" són només "suficients" per a la classificació d'un candidat (el punt a, que és estacionari). Això vol dir el següent: si es dóna la condició, podrem classificar el candidat; si no es dóna, no podrem dir res (doncs hauríem de continuar l'estudi anant a ordres superiors de la fórmula de Taylor).

Esbrinem ara, doncs, un criteri operatiu que ens permeti saber en quin cas estem, atenent només al segon ordre en Taylor. Notem que els desplaçaments des del punt a es poden descriure amb les components del vector $\vec{u}$, $u^i = (x_i - a_i)$. Suposarem de moment el cas senzill de desplaçaments seguint un dels eixos; és a dir, de manera que només varií una coordenada x_j, i per tant totes les components del vector $\vec{u}$ siguin zero tret de $u^j = x_j - a_j$, $\vec{u} = (0, 0, \ldots, u^j, \ldots, 0)$. Per tant, segons Taylor tindrem que, a segon ordre,

$$f(x) \cong f(a) \; + \; \frac{1}{2} H_{jj} \left(u^j\right)^2$$

(recordem que la jota no va sumada en la darrera expressió). Llavors, si la funció ha de decréixer sempre a segon ordre (màxim), serà necessari que tots els elements diagonals d' H siguin negatius, i per a créixer (mínim) tots haurien de ser positius.

Notem ara que la matriu hessiana és simètrica (doncs la continuïtat de les segones derivades parcials implica la igualtat de les creuades, $\partial^2_{ij} f = \partial^2_{ji} f$). Llavors, el teorema espectral ens garanteix que existeix una base ortonormal (és a dir, una reorientació del sistema d'eixos cartesians) en la qual H és diagonal, i els elements de la diagonal en aquesta base reben el nom de valors propis o autovalors de la matriu. Si ens situem en aquest sistema d'eixos, concloem que la condició necessària per al decreixement a segon ordre (màxim) significa que tots els autovalors siguin negatius; la necessària per al creixement (mínim), que tots siguin positius. És fàcil veure que també és suficient: si treballem en aquest sistema d'eixos, un desplaçament qualsevol $\vec{u}$ a partir del punt estacionari a ve descrit a segon ordre, segons Taylor, per

$$f(x) \cong f(a) \; + \; \frac{1}{2} \sum_i H_{ii} \left(u^i\right)^2 ,$$

i per tant a segon ordre sempre "baixem" si els tots els autovalors són negatius (a és màxim) i sempre "pugem" si són positius (a és mínim).

Així doncs, ja tenim un criteri de classificació dels punts estacionaris basat en els autovalors de la matriu hessiana:

· Si a és estacionari i tots els autovalors d' H en a són positius, a és mínim.
· Si a és estacionari i tots els autovalors d' H en a són negatius, a és màxim.
· Si a és estacionari i H té autovalors positius i negatius, llavors a és punt sella.

Com dèiem abans, aquestes condicions són només suficients. Si no se'n dóna cap —quan un o més autovalors són zero—, no podem classificar el candidat (hauríem d'anar a ordres superiors en Taylor).

· CLASSIFICACIÓ de P. ESTACIONARIS amb DETERMINATS (*criteri de Sylvester*):

Existeix un altre mètode de caracterització dels "extrems a segon ordre" equivalent a l'anterior, però més pràctic, perquè no ens obliga a conèixer els autovalors.

L'enunciem fent ús dels "menors principals" de la matrius hessiana, és a dir, dels determinats de les seves submatrius H_m, construïdes agafant les primeres m files i m primeres columnes d' H (sent-hi $m \in \{1,\ 2, \dots, n\}$, que anomenarem ordre del menor).

Aplicant el que es coneix com "criteri de Sylvester de caracterització de formes quadràtiques simètriques", es pot demostrar que:

· Si a és estacionari i $|H_m| > 0$ en a $\forall m$, llavors a és mínim.

· Si a és estacionari i $|H_{senar}| < 0$, $|H_{parell}| > 0$ en a, llavors a és màxim.

· Si a és estacionari i $|H_m| \neq 0$ en a $\forall m$, però no es cap dels casos anteriors, llavors a és un punt sella.

Si un o més dels menors principals s'anul·la, no podem classificar el candidat amb aquest mètode.

NOTA: per a recordar fàcilment aquest criteri, podem situar-nos en el sistema d'eixos on la matriu H és diagonal (i els elements de la diagonal principal són, recordem, els autovalors). Com que sabem que tots els autovalors en un màxim són negatius, automàticament obtenim els signes dels menors principals que diu la taula anterior; anàlogament si es tracta d'un mínim o d'un punt sella.

· CLASSIFICACIÓ de P. ESTACIONARIS en cas de funcions de DUES VARIABLES:

Estudiem la funció $f(x, y)$ en un punt $a(x_0, y_0)$, estacionari d' f.

Per tant, aplicarem el criteri de Sylvester a la següent matriu hessiana:

$$H = \begin{pmatrix} \partial_x^2 f & \partial_{xy}^2 f \\ \partial_{yx}^2 f & \partial_y^2 f \end{pmatrix} = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix},$$

els dos menors principals de la qual són:

$$|H_1| = f_{xx} = \frac{\partial^2 f}{\partial x^2} \qquad \text{i} \qquad |H_2| = |H| = \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2.$$

Conseqüentment,

· Si $|H| > 0$ i $\frac{\partial^2 f}{\partial x^2} > 0$, llavors a és mínim.

· Si $|H| > 0$ i $\frac{\partial^2 f}{\partial x^2} < 0$, llavors a és màxim.

· Si $|H| < 0$, llavors a és punt sella.

NOTES:

1-. En altres casos, per exemple si $|H| = 0$, el criteri no permet afirmar res.

2.- Per una altra banda, pel que fa a la justificació del cas "punt sella" de la taula anterior, notem que en una matriu 2x2 diagonalitzable el determinant és el producte dels seus dos únics autovalors (atès que el determinant no depèn dels eixos usats); per tant si $|H| < 0$ cap dels autovalors pot ser zero i han de tenir signes diferents, i per tant el criteri del signe dels autovalors ens diu que es tracta d'un punt sella.

· EXTREMS CONDICIONATS (mètode dels MULTIPLICADORS de LAGRANGE):

Volem estudiar els extrems de la funció $f(x_1, \dots, x_n)$ sotmesa a les m restriccions expressades per les equacions següents:

$$\begin{cases} g_1(x_1, \dots, x_n) = 0 \\ g_2(x_1, \dots, x_n) = 0 \\ \qquad \vdots \\ g_m(x_1, \dots, x_n) = 0 \end{cases}$$

Definirem la següent "<u>funció de Lagrange</u>", d' $(m + n)$ variables:

$$L(x_1, \dots, x_n;\ \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{i=1}^{m} \lambda_i \cdot g_i(x_1, \dots, x_n),$$

que podem reescriure amb notació més compacta de la forma següent:

$$L(x;\ \lambda) = f + \sum_{i=1}^{m} \lambda_i \cdot g_i$$

Les m noves variables λ reben el nom de "multiplicadors de Lagrange". Si un punt $a(a_1, \dots, a_n)$ és extrem d' f condicionada per les esmentades restriccions, es pot demostrar que llavors existeixen uns valors $b_1, \dots, b_m$ per als multiplicadors que fan que, tot plegat, $(a_1, \dots, a_n;\ b_1, \dots, b_m)$ sigui punt estacionari d' L.

Basant-nos en això, plantegem el "<u>mètode dels multiplicadors de Lagrange</u>", que consisteix en solucionar del següent sistema d' $(m + n)$ equacions:

$$\begin{cases} \dfrac{\partial L}{\partial x_i} = 0, \qquad i = 1, \dots, n \\ \dfrac{\partial L}{\partial \lambda_j} = 0, \qquad j = 1, \dots, m \end{cases}$$

sent-hi $(x;\ \lambda) = (x_1, \dots, x_n;\ \lambda_1, \dots, \lambda_m)$ les incògnites. Si en trobem una solució $(x;\ \lambda) = (a;\ b)$, llavors a és <u>candidat a extrem condicionat</u> d' f. (El mètode no permet esbrinar si, efectivament, els candidats són extrems). Els valors b dels multiplicadors no tenen, en principi, interès en sí mateixos, només els haurem necessitats per a resoldre el sistema.

Una altra forma equivalent del sistema d'equacions del mètode de Lagrange és:

$$\begin{cases} \dfrac{\partial f}{\partial x_i} + \lambda_1 \dfrac{\partial g_1}{\partial x_i} + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_i} = 0, & i = 1, \dots, n \\ g_j(x_1, \dots, x_n) = 0, & j = 1, \dots, m \end{cases}$$

· MULTIPLICADORS de LAGRANGE en el cas 3 VARIABLES i 2 RESTRICCIONS:

Volem estudiar els extrems de la funció $F(x,y,z)$ sotmesa a les restriccions expressades per les equacions següents:

$$\begin{cases} g(x,y,z) = 0 \\ h(x,y,z) = 0 \end{cases}$$

Definim la següent "funció de Lagrange", de 5 variables:

$$L(x,y,z;\ \lambda,\mu) = F + \lambda g + \mu h$$

Les dues noves variables λ i μ són els multiplicadors. Buscarem solucions del sistema format per les 5 equacions següents

$$\begin{cases} \dfrac{\partial L}{\partial x} = 0 & \Rightarrow & \dfrac{\partial F}{\partial x} + \lambda \dfrac{\partial g}{\partial x} + \mu \dfrac{\partial h}{\partial x} = 0 \\ \dfrac{\partial L}{\partial y} = 0 & \Rightarrow & \dfrac{\partial F}{\partial y} + \lambda \dfrac{\partial g}{\partial y} + \mu \dfrac{\partial h}{\partial y} = 0 \\ \dfrac{\partial L}{\partial z} = 0 & \Rightarrow & \dfrac{\partial F}{\partial z} + \lambda \dfrac{\partial g}{\partial z} + \mu \dfrac{\partial h}{\partial z} = 0 \\ \dfrac{\partial L}{\partial \lambda} = 0 & \Rightarrow & g = 0 \\ \dfrac{\partial L}{\partial \mu} = 0 & \Rightarrow & h = 0 \end{cases}$$

on les 5 incògnites són $(x,y,z;\ \lambda,\mu)$. Els tres primers valors (x,y,z) de cada solució del sistema seran les coordenades d'un candidat a extrem. (Els valors dels multiplicadors en cada solució no ens interessen en sí mateixos, només ens hauran sigut necessaris per a resoldre el sistema). El mètode no ens diu, però, com esbrinar si el candidat és realment un extrem.