Λογισμος ΙΙ, Παραγώγιση συναρτήσεων πολλών μεταβλητών

Α. Ν. Γιαννακόπουλος
Τμήμα Στατιστικής

Ο.Π.Α

Εαρινό Εξάμηνο 2018

Μερική παράγωγος

Ας θεωρήσουμε την συνάρτηση $f : U \subset \mathbb{R}^d \to \mathbb{R}$,

$$x = (x_1, x_2, \cdots, x_d) \mapsto f(x_1, x_2, \cdots, x_d).$$

Εστω

$$x_0 = (x_{0,1}, \cdots, x_{0,d}) \in U,$$

ένα σημείο του $\mathbb{R}^d$ και

$$e_i = (\delta_{ij})_{j=1,\cdots,d} = (0, \cdots, 1 \cdots 0),$$

το διάνυσμα που έχει 0 σε όλες τις θέσεις εκτός απο την i θέση.

Ορισμός

Η ποσότητα

$$\frac{\partial f}{\partial x_i}(x_0) = \lim_{\epsilon \to 0} \frac{f(x_0 + \epsilon\, e_i) - f(x_0)}{\epsilon},$$

ονομάζεται η μερική παράγωγος της f *ως προς την μεταβλητη* x_i *στο σημείο* $x_0 = (x_{0,1}, x_{0,2}, \cdots, x_{0,d})$.

Αν ορίσουμε την συνάρτηση $\phi_i : \mathbb{R} \to \mathbb{R}$ με τυπο

$$\phi_i(x_i) = f(x_{0,1}, \cdots, x_{0,i-1}, x_i, x_{0,i+1}, \cdots, x_{0,d}),$$

τότε

$$\frac{\partial f}{\partial x_i}(x_0) = \frac{d\phi_i}{dx_i}(x_{0,i}).$$

Για να υπολογίσουμε την μερικη παράγωγο $\frac{\partial f}{\partial x_i}$ σε κάποιο σημείο αρκεί να θέσουμε όλες τις υπολοιπές μεταβλητές εκτος απο την μεταβλητή x_i σταθερές και ισες με $x_j = x_{0,j}$, $j \neq i$, και μετά να παραγωγίσουμε ως προς την μεταβλητή x_i και τέλος να θέσουμε στο τελικό αποτέλεσμα $x_i = x_{0,i}$.

Παράδειγμα

Αν $f : \mathbb{R}^d \to \mathbb{R}$ *με τύπο*

$$f(x_1, \cdots, x_d) = \exp\Big(-\sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij}x_i x_j\Big),$$

υπολογίστε την μερική παράγωγο $\frac{\partial f}{\partial x_k}$

Ορισμός

Έστω $f : \mathbb{R}^d \to \mathbb{R}$ *μια συνάρτηση. Θα λέμε ότι είναι* **διαφορίσιμη** *στο σημείο* $x_0 = (x_{0,1}, \cdots x_{0,d})^T$ *αν οι μερικές παράγωγοι* $\frac{\partial f}{\partial x_i}$, $i = 1, 2, \cdots, n$ *υπάρχουν στο* x_0 *και*

$$\lim_{x \to x_0} \frac{\left|f(x) - f(x_0) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i})\right|}{\|x - x_0\|} = 0.$$

Αν η συναρτηση $f : \mathbb{R}^d \to \mathbb{R}$ είναι διαφορίσιμη στο σημείο $x_0 = (x_{0,1}, \cdots x_{0,d})^T \in \mathbb{R}^d$ ισχυει ότι για x αρκετά κοντά στο x_0,

$$f(x) \simeq f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i})$$

δηλαδη η γραμμική συνάρτηση $\psi : \mathbb{R}^d \to \mathbb{R}$ με

$$\psi(x) = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i})$$

είναι μια καλή προσέγγιση της f στο σημείο x.

Ορισμός

Έστω $f : \mathbb{R}^d \to \mathbb{R}^n$ μια συνάρτηση. Θα λέμε ότι είναι ***διαφορίσιμη*** *στο σημείο $x_0 = (x_{0,1}, \cdots x_{0,d})^T$ αν οι μερικές παράγωγοι $\frac{\partial f}{\partial x_i}$, $i = 1, 2, \cdots, n$ υπάρχουν στο x_0 και*

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - Df(x_0)(x - x_0)\|}{\|x - x_0\|} = 0,$$

όπου

$$Df(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_d}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x_0) & \cdots & \frac{\partial f_n}{\partial x_d}(x_0) \end{pmatrix}$$

Αν η συναρτηση $f : \mathbb{R}^d \to \mathbb{R}^n$ είναι διαφορίσιμη στο σημείο $x_0 = (x_{0,1}, \cdots x_{0,d})^T \in \mathbb{R}^d$ ισχυει ότι για x αρκετά κοντά στο x_0,

$$f(x) \simeq f(x_0) + Df(x_0)(x - x_0),$$

δηλαδη η γραμμική συνάρτηση $\psi : \mathbb{R}^d \to \mathbb{R}^n$ με

$$\psi(x) = f(x_0) + Df(x_0)(x - x_0)$$

είναι μια καλή προσέγγιση της f στο σημείο x.

Ορισμός

Στην ειδική περιπτωση $f : \mathbb{R}^d \to \mathbb{R}$, ο πίνακας

$$Df(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \cdots, \frac{\partial f}{\partial x_d}(x_0) \right),$$

*ονομάζεται η **βαθμίδα (gradient)** της f στο σημείο $x_0 \in \mathbb{R}^d$ και πολλές φορές χρησιμοποιείται ο συμβολισμός ∇f..*

Στην περίπτωση αυτή ο όρος $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i})$ μπορεί να ερμηνευθεί και ως

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)(x_i - x_{0,i}) = \langle Df(x_0), x - x_0 \rangle,$$

αν θεωρήσουμε το διάνυσμα $x - x_0 = (x_1 - x_{0,1}, \cdots, x_d - x_{0,d})$.

Θεώρημα

Έστω $U \subset \mathbb{R}^d$ *ανοιχτό υποσύνολο του* $\mathbb{R}^d$, *και* $f : \mathbb{R}^d \to \mathbb{R}^n$ *διαφορίσιμη στο* $x_0 \in U$.

Τότε η f *είναι συνεχής στο* x_0.

Διαφορίσιμη σε ανοιχτό σύνολο που περιέχει το $x_0 \implies$ συνέχεια στο x_0.

Προσοχή! Δεν φτάνει όμως μονο να εχουμε τις μερικές παραγώγους σε ένα σημείο για να εξασφαλιστεί η συνέχεια

Παράδειγμα

Η συνάρτηση $f : \mathbb{R}^2 \to \mathbb{R}$ *με τύπο*

$$f(x_1, x_2) = \begin{cases} 1 & \text{αν } x_1 = 0 \text{ ή } x_2 = 0 \\ 0 & \text{αλλιώς} \end{cases}$$

έχει μερικες παραγώγους στο $x_0 = (0,0)$ *εφόσον*

$$\frac{\partial f}{\partial x_1}(0,0) = \frac{\partial f}{\partial x_2}(0,0) = 0,$$

αλλά δεν είναι συνεχής στο $(0,0)$.

Θεώρημα

Έστω $f : U \subset \mathbb{R}^d \to \mathbb{R}^n$ *και οι μερικες παράγωγοι* $\frac{\partial f_i}{\partial x_j}$ *υπάρχουν και είναι συνεχείς σε μια γειτονιά ενός σημείου* $x_0 \in U$.

Τότε η f *είναι διαφορίσιμη στο* x_0.

Ορισμός

Μία συνάρτηση $f : \mathbb{R}^d \to \mathbb{R}^n$ με συνεχείς μερικες παραγώγους θα ονομάζεται C^1 συνάρτηση.

Μια C^1 συνάρτηση είναι και συνεχής.

Ιδιότητες της παραγώγου

Θεώρημα

Έστω $f, g : U \subset \mathbb{R}^d \to \mathbb{R}^n$ *διαφορίσιμες στο* $x_0 \in U$.

1. *Αν* $h = \lambda f$ *όπου* $\lambda \in \mathbb{R}$, *τοτε*
$$Dh(x_0) = \lambda\, Df(x_0).$$
2. *Αν* $h = f + g$ *τότε*
$$Dh(x_0) = Df(x_0) + Dg(x_0).$$
3. *Αν* $n = 1$ *και* $h = fg$ *τότε*
$$Dh(x_0) = g(x_0)Df(x_0) + f(x_0)Dg(x_0).$$
4. *Αν* $n = 1$ *και* $h = \frac{f}{g}$ *τότε*
$$Dh(x_0) = \frac{g(x_0)Df(x_0) - f(x_0)Dg(x_0)}{(g(x_0))^2}.$$

Ο κανόνας της αλυσίδας

Θεώρημα

Έστω οι συναρτήσεις $g : U \subset \mathbb{R}^n \to \mathbb{R}^m$ *και* $f : V \subset \mathbb{R}^m \to \mathbb{R}^p$, *όπου* U, V *ανοιχτά σύνολα.*

Αν η g *είναι διαφορίσιμη στο* x_0 *και η* f *στο* $z_0 = g(x_0)$ *τότε η* $f \circ g$ *είναι διαφορίσιμη στο* x_0 *και*

$$D(f \circ g)(x_0) = Df(z_0)Dg(x_0),$$

όπου το γινόμενο ερμηνεύεται ως γινόμενο πινάκων.

Παράδειγμα (Σύνθεση καμπύλης με συνάρτηση)

Εστω $c = (c_1, \cdots, c_d) : \mathbb{R} \to \mathbb{R}^d$ *μια* d*-διάστατη καμπύλη και* $f : \mathbb{R}^d \to \mathbb{R}$ *μια συναρτηση.*

Αν $h = f \circ c$*, τότε*

$$\frac{dh}{dt}(t) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(c(t))\frac{dc_i}{dt}(t) = \langle \nabla f(c(t)), \frac{dc}{dt}(t)\rangle,$$

όπου

$$\frac{dc}{dt} = \left(\frac{dc_1}{dt}(t), \cdots, \frac{dc_d}{dt}(t)\right).$$

Παράδειγμα

Ποιά η μεταβολή της συνάρτησης $f(x_1, x_2) = \exp(-(x_1^2 + x_2^2))$ ***κατα*** *μήκος της καμπύλης* $x(t) = (cos(t), sin(t))$;

Παράδειγμα

Έστω $f : \mathbb{R}^d \to \mathbb{R}$ *και* $g : \mathbb{R}^d \to \mathbb{R}^d$.

Μπορούμε να εκφράσουμε την $g = (g_1, \cdots, g_d)$ *με*

$$g_i(x) = g_i(x_1, \cdots, x_d), \;\; i = 1, \cdots, d.$$

Αν $h = f \circ g$ *με*

$$h(x_1, \cdots, x_d) = f(g_1(x_1, \cdots, x_d), \cdots, g_d(x_1, \cdots, x_d)),$$

τότε

$$\left(\frac{\partial h}{\partial x_1}, \cdots, \frac{\partial h}{\partial x_d}\right) = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_d}\right) \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_d} \\ \vdots & & \vdots \\ \frac{\partial g_d}{\partial x_1} & \cdots & \frac{\partial g_d}{\partial x_d} \end{pmatrix}$$

Παράδειγμα

Ας πάρουμε μια συναρτηση $f : \mathbb{R}^2 \to \mathbb{R}$ *και ας την δούμε σαν συνάρτηση των καινούργιων μεταβλητών* (r, θ) *οπου* $x_1 = r\cos\theta$ *και* $x_2 = r\sin\theta$.

Υπολογίστε τις μερικες παραγώγους $\frac{\partial f}{\partial r}$ *και* $\frac{\partial f}{\partial \theta}$.

Κάντε μια εφαρμογή στην συνάρτηση $f(x_1, x_2) = \exp(-5(x_1^2 + x_2^2))$.

Ορισμός

Έστω $f : \mathbb{R}^d \to \mathbb{R}$ *και* $h \in \mathbb{R}^d$ *ένα μοναδιαίο διάνυσμα (δηλαδη* $\|h\| = 1$*).*

Η κατευθυνόμενη παράγωγος της f *στο* x *στη διευθυνση* h *ορίζεται ως*

$$(D_n f)(x) = \frac{d}{dt} f(x + t\,n)\Big|_{t=0} = \lim_{t \to 0} \frac{f(x + t\,n) - f(x)}{t},$$

αν το όριο υπάρχει.

Θεώρημα

Ισχυει ότι

$$(D_n f)(x) = \langle Df(x), n \rangle = \langle \nabla f(x), n \rangle = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(x)\, n_i .$$

Η διευθυνση $Df(x) = \nabla f(x)$ είναι η διεύθυνση μεγαλύτερης αύξησης της τιμής της συνάρτησης f.

Για να το δείτε αυτό προσπαθήσετε να λύσετε το πρόβλημα βελτιστοποίησης

$$\max_{n \in \mathbb{R}^d,\ \|n\|=1} (D_n f)(x).$$

Λίγη γεωμετρία δεν βλάπτει ...

Είδαμε ότι μια συναρτηση $c : I \subset \mathbb{R} \to \mathbb{R}^d$ μπορούμε να την κατανοήσουμε ως μια καμπύλη C στον χώρο $\mathbb{R}^d$, υπο την έννοια ότι

$$C = \{(x_1, \cdots, x_d) \in \mathbb{R}^d \ : \ \exists\, t \in I \text{ τ. ω } x_i = c_i(t),\ i = 1, \cdots, d\}.$$

Η συνάρτηση c δίνει την παραμετρική μορφή της καμπύλης C.

Παράδειγμα

Η συνάρτηση $c : [0, 2\pi] \subset \mathbb{R} \to \mathbb{R}^2$ *με τυπο*
$c(t) = (c_1(t), c_2(t)) = (\cos(t), \sin(t))$ *είναι η παραμετρική μορφή ενός κύκλου με κέντρο το σημείο* $(0, 0)$ *και ακτίνα* 1.

Όμοια μια συνάρτηση $f : \mathbb{R}^d \to \mathbb{R}$, μπορούμε να την κατανοήσουμε γεωμετρικά, ως ότι ορίζει μια επιφάνεια S στον χώρο $\mathbb{R}^d$, μέσω της εξίσωσης $f(x_1, \cdots, x_d) = 0$, υπό την έννοια ότι

$$S = \{(x_1, \cdots, x_d) \in \mathbb{R}^d \ : \ f(x_1, \cdots, x_d) = 0\}.$$

Παράδειγμα

Η συναρτηση $f : \mathbb{R}^3 \to \mathbb{R}$ με τύπο $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4$ ορίζει την επιφάνεια μιας σφαίρας στον $\mathbb{R}^3$ με κέντρο $(0, 0, 0)$ και ακτίνα 2.

Παράδειγμα

Η συναρτηση $f : \mathbb{R}^3 \to \mathbb{R}$ με τύπο $f(x_1, x_2, x_3) = \frac{1}{2}x_1^2 + \frac{1}{6}x_2^2 + \frac{1}{10}x_3^2 - 4$ ορίζει την επιφάνεια μιας έλλειψης στον $\mathbb{R}^3$ με κέντρο $(0, 0, 0)$.

Εφαπτομένη μια καμπύλης σε κάποιο σημείο

Ορισμός

Έστω $c : I \subset \mathbb{R} \to \mathbb{R}^d$ *μια διαφορίσιμη καμπύλη και* $x_0 = (c_1(t_0), \cdots, c_d(t_0))$ *ένα σημείο της.*

Το διάνυσμα

$$v := \frac{d}{dt}c(t)\bigg|_{t=t_0} = \left(\frac{d}{dt}c_1(t)\bigg|_{t=t_0}, \cdots, \frac{d}{dt}c_d(t)\bigg|_{t=t_0} \right),$$

ονομάζεται ***εφαπτομένη της καμπύλης*** *στο σημείο* x_0.

Πρόταση

Έστω $f : \mathbb{R}^d \to \mathbb{R}$ *μία* C^1 *συνάρτηση και* $x_0 \in \mathbb{R}^d$ *ένα σημείο επάνω στην επιφάνεια* S *που ορίζεται απο την* $f(x) = 0$.

Το διάνυσμα $Df(x_0) = \nabla f(x_0)$ *είναι κάθετο στην επιφάνεια στο σημείο* $x_0 \in S$ *υπο την ακόλουθη έννοια:*

Αν $c : \mathbb{R} \to \mathbb{R}^d$ *είναι μια οποιαδήποτε καμπύλη* C *επανω στην* S *η οποία περνάει απο το σημείο* x_0 *και* v *ειναι το εφαπτομενο διάνυσμα της* c *στο σημείο* x_0 *τότε*

$$\langle Df(x_0), v \rangle = 0.$$

Απόδειξη

Έστω οποιαδήποτε καμπύλη $c : \mathbb{R} \to \mathbb{R}^d$ επάνω στην S που περνάει απο το x_0.

Άρα $f(c_1(t), \cdots, c_d(t)) = 0$ για κάθε $t \in I \subset \mathbb{R}$, και υπάρχει $t_0 \in I$ τέτοιο ώστε $x_0 = (c_1(t_0), \cdots, c_d(t_0))$.

Η εφαπτομένη της c στο σημείο x_0 δίνεται απο το διάνυσμα

$$v = \left(\frac{d}{dt} c_1(t) \Big|_{t=t_0}, \cdots, \frac{d}{dt} c_d(t) \Big|_{t=t_0} \right).$$

Η συναρτηση $\varphi : I \subset \mathbb{R} \to \mathbb{R}$ που ορίζεται ως $\varphi = f \circ c$ είναι η σταθερή συνάρτηση συνεπώς

$$\frac{d\varphi}{dt}(t) = 0, \ \ \forall \, t \in I.$$

Απο τον κανόνα της αλυσίδας,

$$\frac{d\varphi}{dt}(t) = \langle Df(x(t)), \frac{dc}{dt}(t)\rangle,$$

απο όπου προκύπτει το ζητούμενο.

Ορισμός

Έστω $f : \mathbb{R}^d \to \mathbb{R}$ *μια διαφορίσιμη συνάρτηση η οποία γεωμετρικά μπορεί να κατανοηθεί ως μια επιφάνεια* S *στον χώρο των* d*-διαστάσεων, η οποία ορίζεται απο την εξίσωση* $f(x_1, x_2, \cdots, x_d) = 0$ *και* $x_0 = (x_{0,1}, \cdots, x_{0,d})$ *ένα σημείο της επιφάνειας αυτής (δηλαδή* $f(x_0) = 0$.

Η εξίσωση

$$\langle Df(x_0), x - x_0 \rangle = 0,$$

ορίζει την εξίσωση ενός επιπέδου που είναι το εφαπτόμενο επίπεδο στην επιφάνεια S *στο σημείο* x_0.

Για να το δείτε αυτό αρκεί να παρατηρήσετε ότι αν x ειναι οποιοδήποτε σημείο στο εφαπτόμενο επίπεδο T της επιφάνειας S στο σημείο της $x_0 \in S$ τότε το διάνυσμα $x - x_0 \in T$ (ειναι ένα σημείο του επιπέδου) και συνεπώς θα πρέπει να είναι κάθετο στο κάθετο διάνυσμα στην επιφάνεια S στο σημείο x_0, το οποίο είναι το $Df(x_0)$.

Παράδειγμα

Ένας σκιέρ είναι σε ένα βουνό που η επιφάνεια του περιγράφεται απο την συνάρτηση $f(x_1, x_2, x_3) = 4/10x_1^2 + 3/10x_2^2 + x_3 - 100$ *μεσω της εξίσωσης* $f(x_1, x_2, x_3) = 0$.

Αν ξεκινάει στο σημείο με συντεταγμένες $x_1 = x_2 = 1$, *σε ποιά κατεύθυνση θα πρέπει να αρχίσει να κινείται έτσι ώστε να έχει την πιο γρήγορη διεύθυνση κατάβασης;*

Παράδειγμα

Υπολογίστε τα εφαπτόμενα επίπεδα σε οποιοδηποτε σημείο μιας σφαίρας ακτίνας 1 με κέντρο την αρχή των αξόνων.

Παράγωγοι ανωτερης τάξης

Αν $f : \mathbb{R}^d \to \mathbb{R}$ μια παράγωγίσιμη συνάρτηση, μπορούμε να ορίσουμε τις συναρτήσεις

$$\phi_i : \mathbb{R}^d \to \mathbb{R}, \ \ \phi_i(x) = \frac{\partial f}{\partial x_i}(x), \ \ \forall\, x \in \mathbb{R}^d, \ \ i = 1, \cdots, d.$$

οι οποίες είναι και συνεχείς συναρτήσεις.

Αν οι συναρτήσεις ϕ_i έχουν μερικές παραγώγους τότε μπορούμε να ορίσουμε τις ποσότητες

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \frac{\partial \phi_i}{\partial x_j}(x) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x)$$

Αν οι ποσότητες αυτές είναι επίσης συνεχείς συναρτήσεις θα λέμε ότι η f είναι C^2 συνάρτηση.

Ο ποσότητες $\frac{\partial^2 f}{\partial x_j \partial x_i}$ είναι συνολικά d^2 το πλήθος και μπορεί να χρησιμοποιηθούν για να ορίσουνε σε κάθε σημείο $x \in \mathbb{R}^d$ ένα πίνακα $d \times d$ το πίνακα

$$D^2 f(x) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(x) \right)_{i,j=1,\cdots,d},$$

ο οποίος ονομάζεται πίνακας του Hess.

Θεώρημα (Euler 1734)

Αν $f \in C^2$ *τότε*

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, \cdots, d.$$

Ο πίνακας του Hess για μία C^2 συνάρτηση είναι συμμετρικός.

Απόδειξη

Θα δείξουμε οτι ισχυει για 2 μεταβλητές για να απλουστεύσουμε την απόδειξη.

Έστω

$$I := f(x_{0,1} + \Delta x_1, x_{0,2} + \Delta x_2) - f(x_{0,1} + \Delta x_1, x_{0,2}) \\ - f(x_{0,1}, x_{0,2} + \Delta x_2) + f(x_{0,1}, x_{0,2})$$

Κρατώντας τα $x_{0,2}$, Δx_2 σταθερά, ορίζουμε την συνάρτηση

$$\psi_1(x_1) = f(x_1, x_{0,2} + \Delta x_2) - f(x_1, x_{0,2}),$$

για την οποία ισχύει

$$I = \psi_1(x_{0,1} + \Delta x_1) - \psi_1(x_{0,1}).$$

Κρατώντας τα $x_{0,1}$, Δx_1 σταθερά, ορίζουμε την συνάρτηση

$$\psi_2(x_2) = f(x_{0,1} + \Delta x_1, x_2) - f(x_{0,1}, x_2),$$

για την οποία ισχύει

$$I = \psi_2(x_{0,2} + \Delta x_2) - \psi_2(x_{0,2}).$$

Η συναρτηση ψ_1 είναι μια διαφορίσιμη (παραγωγίσιμη) συναρτηση μιας μεταβλητής οπότε χρησιμοποιώντας το θεώρημα μέσης τιμής για συναρτήσεις μιας μεταβλητής, υπάρχει $\bar{x}_1 \in [x_{0,1}, x_{0,1} + \Delta x_1]$ τέτοιο ώστε

$$\begin{aligned} I &= \psi_1(x_{0,1} + \Delta x_1) - \psi_1(x_{0,1}) = \frac{d\psi_1}{dx}(\bar{x}_1)\Delta x_1 \\ &= \left(\frac{\partial f}{\partial x_1}(\bar{x}_1, x_{0,2} + \Delta x_2) - \frac{\partial f}{\partial x_1}(\bar{x}_1, x_{0,2}) \right) \Delta x_1. \end{aligned}$$

Ορίζοντας τώρα την συνάρτηση μίας μεταβλητής

$$\phi(x_2) = \frac{\partial f}{\partial x_1}(\bar{x}_1, x_2),$$

η παραπάνω γράφεται

$$I = (\phi(x_{0,2} + \Delta x_2) - \phi(x_{0,2}))\, \Delta x_1,$$

Εφαρμοζοντας ξανά το θεώρημα της μέσης τιμής για την ϕ έχουμε

$$
\begin{aligned}
I &= \frac{d\phi}{dx_2}(\bar{x}_2)\Delta x_2 \Delta x_1 \\
&= \frac{\partial^2}{\partial x_2 \partial x_1}(\bar{x}_1, \bar{x}_2)\Delta x_2 \Delta x_1,
\end{aligned}
$$

οπότε

$$
\frac{\partial^2}{\partial x_2 \partial x_1}(\bar{x}_1, \bar{x}_2) = \frac{1}{\Delta x_1 \Delta x_2} I,
$$

και απο την συνέχεια των δευτερων μερικών παραγώγων,

$$
\frac{\partial^2}{\partial x_2 \partial x_1}(x_{0,1}, x_{0,2}) = \lim_{(\Delta x_1, \Delta x_2) \to (0,0)} \frac{1}{\Delta x_1 \Delta x_2} I.
$$

Με όμοιο τρόπο η ψ_2 είναι συνεχής συνάρτηση και εφαρμοζοντας το θεώρημα μεσης τιμής 2 φορές οπως και παραπάνω καταλήγουμε ότι

$$\frac{\partial^2}{\partial x_1 \partial x_2}(x_{0,1}, x_{0,2}) = \lim_{(\Delta x_1, \Delta x_2) \to (0,0)} \frac{1}{\Delta x_1 \Delta x_2} I.$$

Απο την μοναδικότητα του ορίου προκύπτει το ζητούμενο.

Σειρές Taylor

Ας υποθέσουμε ότι εχουμε μια συνάρτηση $f : \mathbb{R}^d \to \mathbb{R}$ η οποία είναι αρκετές φορές συνεχώς παραγωγίσιμη στο σημείο $x_0 \in \mathbb{R}^d$.

Η συνάρτηση $\phi : \mathbb{R} \to \mathbb{R}$ η οποία ορίζεται ως

$$\phi(t) = f(x_0 + t\, z)$$

είναι παραγωγίσιμη στο σημείο $t = 0$ για οποιαδήποτε επιλογή του διανύσματος $z \in \mathbb{R}^d$.

Εφαρμοζουμε τώρα το ανάπτυγμα Taylor για την συναρτηση $\phi : \mathbb{R} \to \mathbb{R}$, συμφωνα με το οποίο

$$\phi(t) = \phi(0) + \left.\frac{d\phi}{dt}\right|_{t=0} t + \frac{1}{2}\left.\frac{d^2\phi}{dt^2}\right|_{t=0} t^2 + \cdots + \frac{1}{n!}\left.\frac{d^n\phi}{dt^n}\right|_{t=0} t^n + O(t^n)$$

Απο την συνέχεια της f,

$$\phi(0) = f(x_0).$$

Απο τον κανόνα της αλυσίδας και την συνέχεια

$$\frac{d\phi}{dt}(t) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(x_0 + t\,z) z_i = \langle Df(x_0 + t\,z), z \rangle,$$

οπότε

$$\frac{d\phi}{dt}(0) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(x_0) z_i = \langle Df(x_0), z \rangle.$$

Εφαρμόζοντας ξανά τον κανόνα της αλυσίδας

$$\begin{aligned}\frac{d^2\phi}{dt^2}(t) &= \frac{d}{dt}\left(\sum_{i=1}^{d}\frac{\partial f}{\partial x_i}(x_0 + t\,z)z_i\right)\\ &= \sum_{i=1}^{d}\frac{d}{dt}\left(\frac{\partial f}{\partial x_i}(x_0 + t\,z)\right)z_i\\ &= \sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}(x_0 + t\,z)\right)z_j z_i\\ &= \sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\partial^2 f}{\partial x_j \partial x_i}(x_0 + t\,z)z_j z_i,\end{aligned}$$

και χρησιμοποιώντας την ιδιότητα της συνέχειας

$$\frac{d^2\phi}{dt^2}(0) = \sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\partial^2 f}{\partial x_j \partial x_i}(x_0)z_j z_i$$

Μπορούμε να συνεχίσουμε όμοια για την τρίτη παράγωγο

$$\frac{d^3\phi}{dt^3}(0) = \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(x_0) z_k z_j z_i,$$

και εν γένει

$$\frac{d^n\phi}{dt^n}(0) = \sum_{i_1=1}^{d}\cdots\sum_{i_n=1}^{d} \frac{\partial^n f}{\partial x_{i_n} \cdots \partial x_{i_1}}(x_0) z_{i_1} \cdots z_{i_n},$$

Μαζευοντας ολα τα παραπάνω καταλήγουμε στον τύπο του Taylor για συναρτήσεις d μεταβλητών

$$f(x_0 + tz) = f(x_0) + \sum_{i_1=1}^{d} \frac{\partial f}{\partial x_{i_1}}(x_0) z_{i_1} + \frac{1}{2} \sum_{i_1=1}^{d} \sum_{i_2=1}^{d} \frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}} f(x_0) z_{i_1} z_{i_2} +$$

$$\cdots + \frac{1}{n!} \sum_{i_1=1}^{d} \cdots \sum_{i_n=1}^{d} \frac{\partial^n f}{\partial x_{i_n} \cdots \partial x_{i_1}} f(x_0) z_{i_1} \cdots z_{i_n} + O(\|z\|^n).$$