**Mínim nombre de computacions per a un determinat terme del polinomi de Taylor**

(Cas especial: 2 variables)

Autor: Junhao Liu

Agraïment especial als professors que m’han guiat en la demostració, especialment la Montserrat Maureso Sánchez, en José Luís Ruiz Muñoz i la Anna De Mier Vinué.

INTRODUCCIÓ

Per a un polinomi de Taylor d’una sola variable, el nombre de vegades que cal derivar sobre una funció per a obtenir un terme específic és evident (és el grau del polinomi). És a dir, per exemple, suposem una funció , de la qual es vol calcular el polinomi de Taylor de grau centrat en el punt , es té un polinomi d’aquesta forma:

Cosa que també es pot expressar de forma:

Llavors denotem cada terme del sumatori com “terme” del polinomi de Taylor.

És evident que, per a obtenir el primer terme del polinomi (), cal derivar la funció 0 vegades, per tal que a continuació s’apliqui a , dividint el resultat per , etc. De forma anàloga, per al segon terme (), es deriva 1 vegada la , per al tercer terme (), es deriva 2 vegades la , i així successivament. Llavors, per a obtenir la funció d’un cert terme, cal derivar la vegades.

Ara bé, per al polinomi de Taylor de 2 variables, la cosa es complica. Per a un “terme” del polinomi, es cal uns “subtermes” del terme. Sigui un polinomi de Taylor de grau de la funció centrat en , l’expressió del polinomi és així (suposant que la compleix les condicions del teorema de Schwarz, és a dir, . Aquesta suposició és molt important perquè l’expressió següent sigui certa, i conseqüentment que el meu plantejament tingui sentit):

Per exemple, en un , el “subterme” per a :

Aquí, i similarment per a qualsevol altre , el pas més “complex” computacionalment és les derivades parcials.

Sembla que per a la primera funció, cal derivar 4 vegades sobre , per a la segona funció, cal derivar 4 vegades, 3 sobre i 1 sobre , així successivament. Resultant que cal derivar en total vegades. Malgrat això, quan es calcula , s’ha de primer passar per la , després per la , i , finalment obtenint el resultat desitjat. Però, llavors, per a calcular el , **no cal** derivar 4 vegades des de , sinó que es pot **aprofitar** el resultat intermedi i aplicar només 1 pas sobre obtenint la que, segons la suposició del teorema de Schwarz, és la mateixa funció que . Anàlogament per als altres, i, veiem que excepte el , tots podem aprofitar un resultat intermedi, així no havent de derivar vegades per a totes elles. Simètricament, per al on es deriva més vegades sobre que , podem aprofitar els resultats intermedis de per a estalviar encara més computacions.

Podem fer un petit gràfic del màxim estalvi que es pot arribar a tenir:

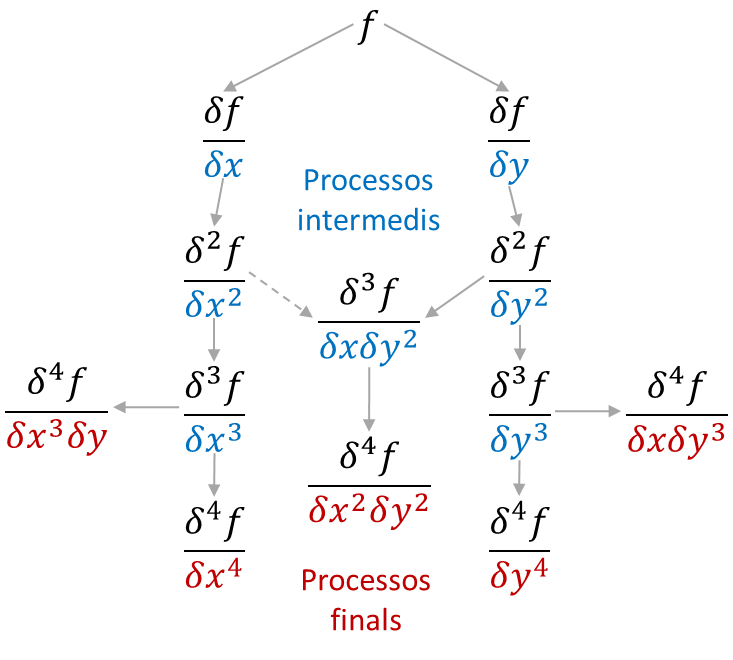


Fig. I.

Veiem que, per a , no cal derivar en total 20 vegades, sinó amb només 12 vegades ja és prou perquè s’ha pogut aprofitar els processos intermedis.

PLANTEJAMENT

Per a un “subterme” qualsevol d’un polinomi de Taylor de funció de 2 variables, quin és el mínim nombre de derivacions que cal computar, aprofitant tots els passos intermedis, per a obtenir totes les funcions derivades necessàries del “subterme”?

(si es vol profunditzar més, la hipòtesi es podrà generalitzar per a funcions de nombre de variables qualsevol)

EXPERIMENTACIONS

Vaig fer el mateix gràfic (similar a la fig. I) per a més ’s. Es va obtenir els resultats següents:

|  |  |  |
| --- | --- | --- |
| Grau de polinomi | “Vegades (més òptimament)” a derivar | Esquema visual |
| 1 | 2 |  |
| 2 | 5 |  |
| 3 | 8 |  |
| 4 | 12 |  |
| 5 | 16 |  |
| 6 | 20 |  |

TESI

El nombre de derivacions òptim per a un subterme de grau s’aproxima a .

DEMOSTRACIÓ

Aquí demostrarem per a casos especials, on és una potència de 2, per tant, dona un resultat exacte, cosa que, segons la meva demostració a continuació, serà certa i exacte. A partir d’això es demostrarà la tesi fàcilment.

La idea principal és, com que la és de 2 variables, només hi ha 2 vies per a obtenir la següent derivada: (1) derivar 1 vegada sobre , (2) derivar 1 vegada sobre . Podem transformar el problema en punts en un espai de coordenades , on l’eix d’abscisses representa “les derivades sobre ”, i similarment, l’eix d’ordenades representa “les derivades sobre ”. Més concretament, la posició correspon amb la funció original sense derivar, “derivar 1 vegada sobre ” significa “moure 1 pas cap a la direcció i sentit ”, “derivar 1 vegada sobre ” significa “moure 1 pas cap a la direcció i sentit ”. Aquí no considerem els moviments amb vector ni , tot i que, segons el Teorema Fonamental de Càlcul, podem assolir-lo mitjançant integracions, però, es perd la informació de la constant (el de les primitives).

Llavors, per exemple, per al subterme , aquestes són totes les derivades possibles:

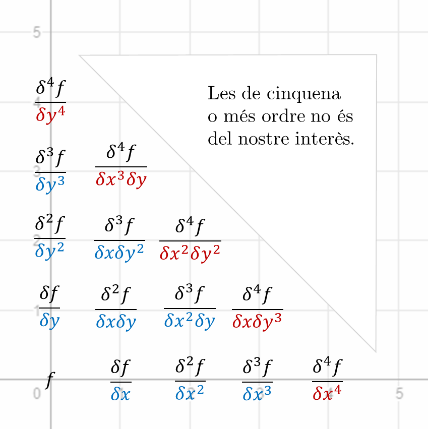


Fig. II. Esquema per a .

On les derivades en blau són tots els possibles “resultats intermedis”, i les en roig són els “resultats finalitat”. Es vol construir camins des d’ fins a cadascú dels resultats finalitat, aprofitant el màxim els altres camins. El següent seria un exemple:

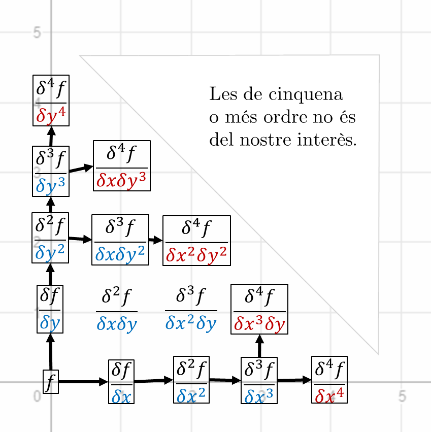


Fig. III. Camí 1 per a .

Observant aquest esquema (fig. III) veiem que, per a obtenir totes les derivades de quart grau, es pot estalviar passar per o . Comptem el nombre de “fletxes”, que és certament el nombre de derivacions fetes, i observem que és 12. Observem que té una estructura d’un arbre binari, és a dir, per a estalviar al màxim, totes les funcions només pot “rebre” com a molt 1 fletxa (les enquadrades), i en casa contrari 0 (les no enquadrades). I per a maximitzar la utilitat d’una derivada parcial intermèdia, ha de “apuntar” a 2 derivades de següent grau, o com a mínim 1.

També és curiós veure que, totes les derivades “finalitat” se situen en punts on , formant-se una recta. És evident que això es justifica pel fet que totes les derivades finalitat són del quart ordre ().

A continuació teniu un esquema per a (fig. IV):

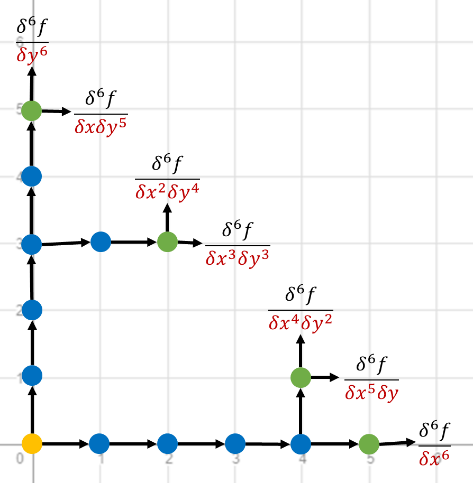


Fig. IV. (els resultats intermedis s’han abstret en “nodes”)

De cara a la demostració, pensem de manera recursiva a partir dels nodes finalitat. Abans vam comentar que, per a aprofitar al màxim la utilitat d’un node, d’aquest node cal “sortir 2 fletxes”, una amb direcció i l’altra amb direcció . Llavors suposem un node d’aquests per a cada parell dels nodes finalitat situats a la recta :

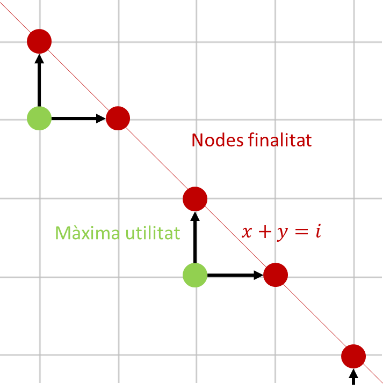


Fig. V. Nodes de màxima utilitat (en verd).

De forma anàloga, situem nodes als que cal el mínim nombre de “fletxes” per a connectar cada dues nodes de màxima utilitat. Aquests nodes han d’estar obligatòriament en posicions on . L’únic punt que satisfà aquesta condició distant el mínim als dos nodes de màxima utilitat és:

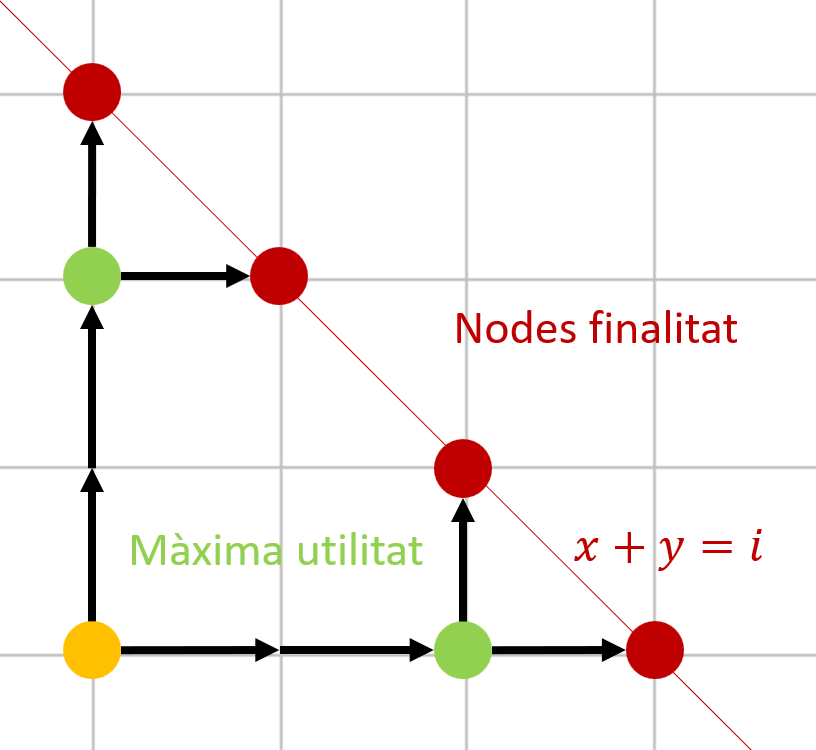


Fig. VI. Nodes “connectors” (en taronja).

Comencem veient un patró. Quan vam situar els nodes de màxima utilitat, vam ajuntar els nodes finalitat en “conjunts” de 2 “fletxes”, units pel node de màxima utilitat. Després, vam adjuntar dos conjunts d’aquests amb un node nou, formant un conjunt de “fletxes” més gran. Observem que el conjunt petit correspon amb l’esquema per a (fig. VII) i aquest gran correspon amb l’esquema per a (fig. VIII):

(s’ha abstret de nou, les fletxes es representen en segments)

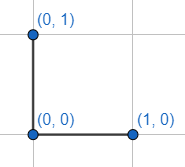
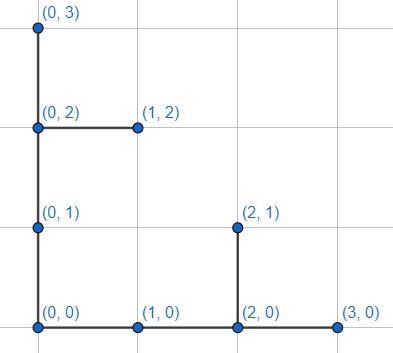
 

Fig. VII. Fig. VIII.

Cal destacar que, al conjunt petit, entre els 2 nodes finalitat i el de màxima utilitat, s’ha “afegit 2 fletxes” (+2 derivacions necessàries totals), i entre 2 conjunts s’ha “afegit 4 fletxes” (+4 derivacions). Apliquem la mateixa lògica descrita fins ara, i per a aprofitar al màxim, situem el node de següent nivell per a ajuntar dos conjunts grans del procés anterior:

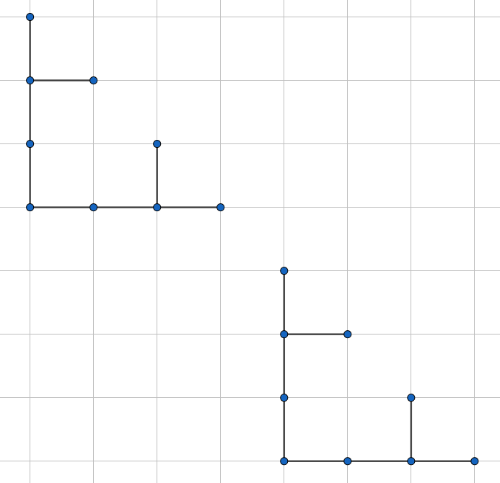
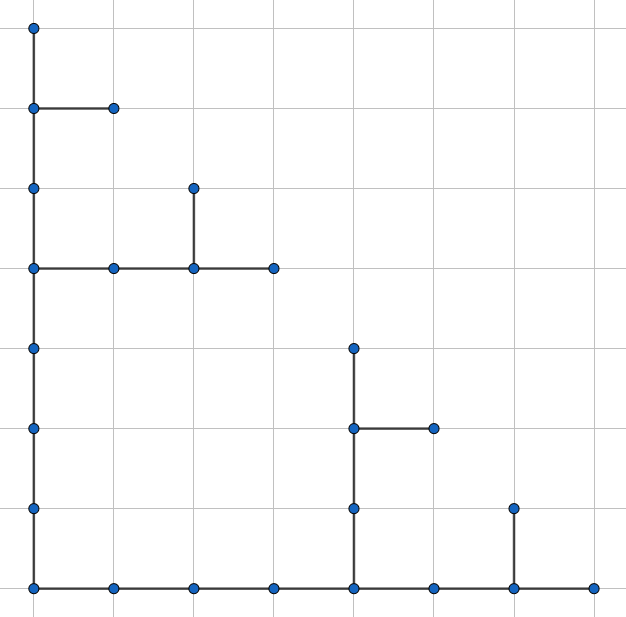
 

Fig. IX. 2 conjunts. Fig. X. 2 conjunts de nodes ajuntats mitjançant un node.

Fixem que, les “fletxes” (ara per les abstraccions, es veu implícit) que cal per a ajuntar els 2 conjunts de menor nivell són 8 (4 fins al conjunt de dalt, i 4 al conjunt de dreta). Anàlogament, per al conjunt del següent nivell, es seguirà la mateixa estructura: l’altre conjunt de nivell actual està situat a la mateixa diagonal, i, per a connectar-los, la posició òptima serà , amb la posició més petita d’entre tots els nodes, i la la més petita també. Per tant, en total es calen fletxes de quantitat diferència d’ “altura” entre els dos conjunts (la diferència de la posició més petita d’un conjunt i la posició més petita de l’altre), més la quantitat de diferència d’ “distància” (la diferència de la posició més petita d’un conjunt i la posició més petita de l’altre), coses que, es veu evident que seran de mateix valor, ja que els dos conjunts se situen de manera perfectament diagonal. En aquest cas concret, el nombre de “fletxes” novament afegits del conjunt del següent nivell és , i tindrà l’estructura anàloga a l’esquema per a .

Veiem que cada vegada el nombre de “fletxes” novament afegits creix exponencialment a base 2, al principi per a ajuntar els 2 nodes “finalitat” calia fletxes, després per a ajuntar dos conjunts calia , després , i , per la mateixa lògica, després vindrà , , ...

Al mateix temps, hem observat que l’estructura d’unir dos nodes “finalitat” és isomorf a l’esquema de fletxes mínim per a (fig. VII), al següent nivell és isomorf a l’esquema per a (fig. VIII), després es veu evident que és isomorf al de per a , i així successivament.

Amb tot lo anterior, ara només ens queda enllaçar els dos conceptes: quina és la relació entre les vegades d’ajuntar conjunts i el seu “isomorfisme” a un esquema de grau ? Si arribem a resoldre aquesta pregunta, la demostració queda acabada. Compte, que aquesta demostració només és exacta quan . Justificarem els altres casos quan es finalitza aquesta.

Per a allò, ara concentrem només en l’eix d’ordenades (també es pot fer amb l’eix abscisses canviant tots els termes, però voldran dir el mateix). Obliguem a què, quan s’aplica l’acció “ajuntar dos conjunts”, només es fa amb el conjunt que es troba a sota d’aquesta. És a dir, fixem l’ d’entre tots els nodes del conjunt en aplicar l’acció “ajuntar”:

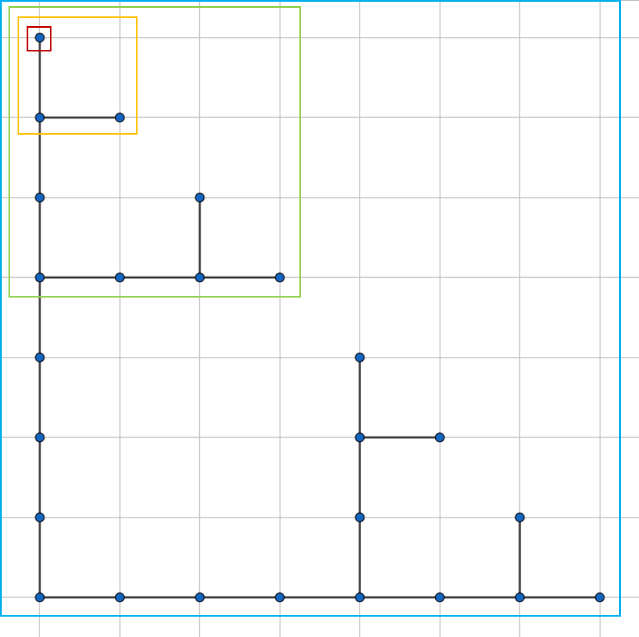
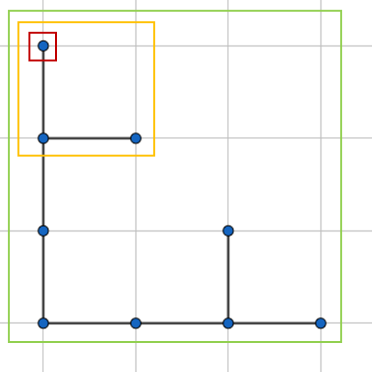
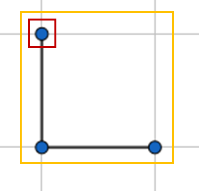


Fig. XI. Els quadrats sempre es mantenen en la mateixa posició després d’ajuntar.

Es pot observar que (fig. XI), la longitud de la composició de “fletxes” (el “pal”, fig. XII) al més esquerra del conjunt és la suma de potències de 2. Per al 4t esquema de la fig. X, per exemple, la longitud és . Si ajuntem una vegada més, lògicament, la longitud queda , que és la suma de “fletxes afegits en ajuntar la 1a vegada” + “fletxes afegits en ajunta la 2a vegada” + “... la 3a vegada” + “... la 4a vegada”. Fixem aquesta composició de “fletxes” (el “pal”) a l’eix de ordenades d’esquema del subterme , i descobrim que, només mirant la longitud d’aquest “pal”, podem dir quantes vegades s’ha aplicat l’acció “ajuntar”, sempre que la seva longitud compleixi :

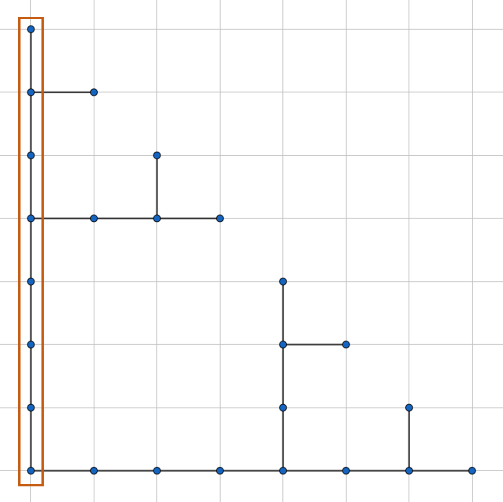


Fig. XII. La “composició de fletxes” referenciada (o el “pal”).

Però, aquest conjunt és anàleg a l’esquema del subterme del polinomi! Llavors, només ens queda trobar la relació entre aquesta longitud i el nombre total de fletxes intervingudes en l’esquema.

Demostració per inducció completa. Sigui el conjunt d’un sol punt el conjunt de nivell 0, i en cada acció d’ “ajuntar”, el nivell del conjunt format puja +1 nivell.

Per a

En efecte, en cada acció d’ “ajuntar”, s’aporta fletxes al nombre total,

UTILITAT

Al final, quina utilitat té aquesta tesi? Primer de tot, a la ciència de les matemàtiques és tot sobre la diversió i bellesa. Dit això, d’entre totes les possibles utilitats, una que és l’única que se m’ha acudit a mi és el resto de Lagrange. Suposem que volem calcular el resto de Lagrange d’un polinomi de Taylor ja calculat, és a dir, on es té el polinomi ja fet i s’ha perdut totes les derivades parcials originals. En aquest cas, amb es pot predir ràpidament quanta computació és mínimament necessària per a calcular-lo.