Nemlineáris egyenletrendszer gyökeinek meghatározása

A nemlineáris egyenletek megoldása lényegesen nehezebb probléma mint a lineáris egyelnetrendszereké. Egy dimenzióban vannak csak hatékony megoldások, több változó esetén nincs olyan általános megoldás amely minden probléma öszzes megoldását hatékonyan megtalálná. Ennek fő oka, hogy míg egydimenzióban a gyökök lokalizálhatók egy intervallumon belül és az intervallum szűkítésével megtalálhatóak, több dimenzióban ez nem tehető meg. A probléma egy dimenzióban az egyenlet átrendezésével így írható fel:

$$f(x) = 0$$

ahol f nemlineáris függvénye x-nek. A több dimenziós probléma azt jelenti, hogy a függvény több x_j $j = 1...M$ változótól függ. Ha eygenletrendszerünk van akkor több $f_i(x_1, ...x_j...x_M)$ $i = 1...N$ függvény gyökét keressük szimultán módon. Általában ha $N > M$ akkor nincs megoldása az egyenletrendszernek, viszont ha $N \leq M$ akkor sem mondható általában, hogy van megoldás, és az sem, hogy egyértelmű megoldás van. Elég gyakori aa a degenerált eset, hogy a gyökök nem is diszkrétek, hanem egy vagy több folytonos intervallumon vannak.

Gyorsabb módszerek vannak abban az esetben ha nem csak a függvényt, de annak deriváltját is ismerjük.

Az egy dimenziós esetet tárgyaljuk részletesen.

A gyökök bekeretezése

(ábrák a lehetséges gyökökről Numrec 9.1.1)

A nemlineáris gyökkereső algoritmusok iterációval érnek célhoz. Az iterációknál nagyon fontos a jó kezdőpont meghatározása. A jó kezdőérték megtalálása kívül esik a numerikus algoritmusok hatáskörén, a probléma analizálásával kaphatunk becsléseket. Ez akkor még fontosabb, ha az egyenletnek több gyöke is van, de mi ezek közül csak egyet (a *fizikai megoldást*) keressük. Jó segítséget nyújthat a függvény grafikonjának megszemlélése abban, hogy sejtésünk legyen arról, milyen gyökökre számíthatunk. Egy dimenzióban a gyök helye elég jól bekeretezhető, hiszen itt előjelet vált a függvény. Kivétel ez alól, amikor több gyök egymáshoz nagyon közel van. Ha páros számú gyök van egymáshoz közel akkor lehet, hogy nagyon kicsi az az intervallum ahol előjelet vált a függvény. Ekkor minimum/maximum kereséssel érhetünk célhoz.

A kezdeti bekeretezés normál esetben megtehető úgy pl., hogy kiindulunk egy intervallumból, és ha nem különbözik a függvény előjele a két határon, akkor azt a határt amelyiken a függvény abszolút értéke kisebb, kitoljuk egy adott faktorral (pl. megduplázzuk arra az intervallumot). Ha több gyökre számítunk akkor a teljes intervallumot feloszthatjuk akkora intervallumokra ami a tipikus gyöktávolságnál kisebb, és kikeressük azokat amelyekben van előjelváltás.

A gyök megkeresése

Felezős (bisection) módszer: A legegyszerűbb eljárás. Miután bekereteztük a gyököt keressük ki az intervallum felező pontját. Azt a végpontot amelyiknél a függvényérték előjele megegyezik a középső pontiéval, lecseréljük a középső pontra. A módszer minden esetben konvergál. Ha több gyök van akkor közülük egyet talál meg, ha nincs gyök csak szakadás vagy szingularitás miatt vált előjelet a függvény, akkor ezt a pontot találja meg.

A módszer minden lépésben felezi az intervallumot, így ha a kezdeti intervallum hossza ϵ_0 volt, és a kívánt pontosság ϵ, akkor $n = \log_2(\epsilon_0/\epsilon)$ lépésben kapunk eredményt. Nyilván ϵ-t a gyök nagyságrendjének megfelelően választjuk meg, vagy pedig (ha nem 0 közeli a gyök) relatív hibahatárt adhatunk meg leállási feltételként.

(ábrák)

Szekáns módszer: Ez a módszer általában gyorsabb mint a felezős módszer. Míg a felezős módszernél az intervallum egy lépésben konstansszorosára (felére) csökkent, vagyis lineárisan változik itt a konvergencia rendje

$$\lim_{k\to\infty} |\epsilon_{k+1}| = const \times |\epsilon_k|^{1.618}$$

magasabb hatvánnyal csökken. A módszer akkor használható, ha a függvény elég sima olyan értelemben, hoyg a vizsgált intervallumban a függvény lineárissal viszonylag jól közelíthető. Kössük össze a függvény két végpontját egy egyenessel. Ahol ez metszi a tengelyt az lesz az egyik új pontunk, és a meglévő két pontból a régebbit hagyjuk el. Az első lépésben praktikus a nagyobb abszolútértékű pontot a régebbinek tekinteni. Ennek a módszernek hátránya, hogy nem tartja bekeretezve a gyököt, bizonyos esetekben el is távolodhat tőle a végtelenbe.

Regula falsi (false position) módszer: Ez a módszer általában gyorsabb mint a felezős módszer, viszont lassab mint a szekáns módszer, de a gyököt bekeretezve tartja. Az algoritmus az előzőhöz hasonló, de a metszéspont megtalálása után nem a legrégebbi pontot hagyjuk el, hanem azt, amelyiknél a függvény előjele megegyezik a metszéspontiéval. Így továbbra is bekeretezve marad a gyök.

Magasabb rendű módszerek: A fenti módszerek kiterjeszthetőek úgy, hogy ahelyett, hogy két pontot tartunk meg minden lépésben és arra egyenest fektetünk, megtartunk több pontot és magasabbrendű görbét interpolálva találjuk meg a következő pontot. Ilyen pl. az inverz kvadratikus interpolációs módszer, ahol $x = ay^2 + by + c$ alakot interpolálunk.

(ábra 9.2.3)

Vannak olyan függvények viszont ahol a szekáns és regula falsi módszerek nagyon lassan konvergálnak lemaradva még a felezős módszertől is. Így a legjobb az, ha kombináljuk a biztos de általában lassabb felezős módszert a gyorsabb de bizonyos függvényeknél rosszul viselkedő pl. inverz kvadratikus módszert. Ez a Wijngaarden-Dekker-Brent (röviden Brent) módszer. Az algoritmus figyeli, hogy az inverz kvadratikus módszerrel kapott új pont belül van-e a kereten, ha nem, vagy az így adódó lépés túl kicsi akkor inkább felezi az intervallumot.

Newton-Raphson módszer: Az eddigi módszerek nem használták a függvény deriváltját, csak magát a függvényt. Abban az esetben ha a függvény deriváltja is adott explicit módon, akkor ennek segítségével gyorsabb módszer dolgozható ki. Az algoritmus nem használ bekerertezést, csak egy pontra van szükségünk, mint kezdeti becslés. Ebben a pontban a derivált segítségével meghúzzuk a függvény érintőjét, és a tengellyel való metszéspont lesz az új közelítés. Ha ugyanis ismerjük a függvényt az $f(x)$ pontban és keressük azt a δ értéket amire $f(x+\delta)=0$ akkor ha a függvény kellően sima

$$f(x+\delta) \approx f(x) + f'(x)\delta + \frac{1}{2}f''(x)\delta^2 + ...$$

elég a Taylor-sorból csak az elsőrendű tagot megtartani és így

$$\delta = -\frac{f(x)}{f'(x)}$$

Könnyen belátható, hogy a konvergencia kvadratikus, vagyis

$$\epsilon_{i+1} = const \times \epsilon_i^2$$

Van sok olyan eset amikor a módszer nem konvergál (ábrák) ezért a Brent módszerhez hasonlóan érdemes hibrid módszert kialakítani a felezős algoritmussal kombinálva.

Ha nem ismert a derivált explicit módon, akkor az egy dimenziós esetben nem érdemes numerikus deriváltat venni. Az egyrészt a további függvénykiértékelés miatt lassítja az algoritmust, másrészt mivel a numerikus derivált kevésbé pontos, a konvergencia rendje tovább csökken, így pl. a szekáns módszer gyorsabb.

A Newton-Raphson módszer alkalmazható lokális gyorsasága de globális sérülékenysége miatt úgy is, hogy más módszerrel jól megközelítjük a gyököt, és a Newton-Raphson módszerrel pedig pontosítjuk.

Polinomok gyökei

Egy n-ed rendű polinomnak n gyöke van. A gyökök lehetnek valósak vagy komplexek, ha az együtthatók valósak, akkor a gyökök vagy valósak, vagy pedig komplex konjugált párok. Előfordulhatnak többszörös gyökök is. Ha a polinom összes gyökére kíváncsiak vagyunk akkor praktikus úgy eljárni, hogy miután egy gyököt (r) megtaláltunk, a polinomot leosztjuk $(x - r)$-el, ami viszonylag eygszrű algoritmussal megtehető. $(P(x) = (x - r)Q(x))$ Az így kapott polinom egyel alacsonyabb rendű, és a gyökkereső algoritmus egy új gyökhöz fog tartani. Mivel a polinom osztások során felhalmozódhatnak a numerikus hibák, célszerű a végén a kapott gyököket pl. Newton-Raphson módszerrel pontosítani. A [1] leírja a Laguerre módszert, amely képes komplex gyökök megkeresésére is.

Többdimenziós gyökkeresés

Általános algoritmus amely mindig konvergál nem létezik. A probléma az, hogy nagyon nehéz behatárolni a gyököt. Vegyünk csak 2 dimenziót és két függvényt. Mindkét függvény metszete a 0 síkkal (előre nem ismert számú) zárt görbékből áll. Ezek közös pontjai az egyenletrendszer gyökei. A fenti kontúrok teljes feltérképezése általában lehetetlen.

Ha azonban tudjuk, hogy hány gyökre számítunk, és jó közelítéssel azt is, hogy hol, akkor az egydimenziós módszerek viszonylag egyszerűen kiterjeszthetőek.

Pl. a Newton-Raphson módszer esetében a Taylor-sor: s

$$f_i(\mathbf{x}+\delta\mathbf{x}) = f_i(\mathbf{x}) + \sum_{j=1}^{N} \frac{\partial f_i}{\partial x_j}\delta x_j + ...$$

Ha bevezetjük a Jacobi mátrixot

$$J_{ij} = \frac{\partial f_i}{\partial x_j}$$

a teljes egyenletrendszer kifejtése első rendig így írható

$$\mathbf{f}(\mathbf{x}+\delta\mathbf{x}) = \mathbf{f}(\mathbf{x}) + \mathbf{J}\cdot\delta\mathbf{x}$$

és hasonlóan mint egy dimenzióban keressük azt a $\delta\mathbf{x}$ vektort amely a gyökhöz lépteti az egyenletrendszert, vagyis 0-vá teszi a bal oldalt. A megoldást a

$$\mathbf{J} \cdot \delta\mathbf{x} = -\mathbf{f}$$

lineáris egyenletrendszer adja. Ezt pl. LU dekompozícióval megoldhatjuk és megkapunk egy többdimenziós Newton-Raphson lépést.