

ALGORITMI GENETICI DE OPTIMIZARE

George Daniel Mateescu

Rezumat. Algoritmii generici reprezintă un instrument util pentru rezolvarea unei clase largi de probleme, pornind de la principii extrase din biologie. Scopul acestui articol este de a pune în evidență principiul achizițiilor genetice, exprimate matematic prin mutații pe direcția gradientului unei funcții obiectiv.

Notății

Ne propunem minimizarea funcției $f : D \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$, în care, formal, putem considera că D reprezintă o mulțime de restricții ale problemei.

Soluții admisibile.

Evident, orice element din D reprezintă o soluție admisibilă a problemei de optimizare. Deoarece un element din D este caracterizat de componentele sale, (x_1, x_2, \dots, x_m) , $x_i \in \mathbb{R}$ vom considera că fiecare x_i este un cromozom. Desigur, inspectarea tuturor elementelor din D nu este posibilă, din cauză că mulțimea este infinită, iar prin comparație cu un model biologic, putem considera că mulțimea tuturor combinațiilor cromozomiale este infinită.

Populație.

Asemănător cu lumea biologică, vom considera că populația este o submulțime a mulțimii soluțiilor admisibile, la fel cum populația vie este o submulțime a tuturor combinațiilor cromozomiale. Evident, populația variază în timp, astfel încât vom nota prin $P(t)$ populația existentă la un anumit moment dat. În cazul nostru, aceasta reprezintă o submulțime finită a mulțimii D . Un individ din $P(t)$ va fi cu atât mai bine adaptat cu cât zestrea sa cromozomială conduce la o valoare mai mică a funcției obiectiv, $f(x_1, x_2, \dots, x_m)$.

Selecție.

Există mai multe metode de simulare a selecției biologice (naturale) care în mod esențial sunt cuprinse între două extreme: selecția elitistă, în care numai cei mai buni supraviețuiesc și selecția pur întâmplătoare. Vom folosi o selecție elitistă care corespunde stării unei populații având resurse limitate. Concret, vom considera că numărul total al indivizilor este constant, reținând la fiecare etapă numai pe cei mai bine adaptați. Din punct de vedere matematic, vom sorta indivizii în funcție de valorile crescătoare ale funcției obiectiv și vom elimina surplusul de populație.

Mutații genetice.

Mutațiile reprezintă mici variații cromozomiale, care pot fi întâmplătoare, afectând o parte a populației. În cazul nostru, mutațiile vor fi reprezentate de mici modificări ale coordonatelor punctelor din populația curentă. Dacă acceptăm posibilitatea unor mutații de tip “*achiziții genetice*”, adică mutații care apar ca urmare a eforturilor indivizilor de adaptare la mediu, atunci putem găsi un corespondent în problema noastră de optimizare, în mici variații “în sensul creșterii

performanței”. Matematic, ar însemna ca micile modificări ale coordonatelor să fie făcute pe direcția gradientului, adică pe acea direcție care modifică funcția obiectiv, în sensul dorit.

Formal, vom considera că $u + h$ este o mutație genetică a lui u , pentru h suficient de mic. Într+adevăr, pornind de la formula lui Lagrange:

$$f(u + h) = f(u) + f'(u + \tau h)h, \quad \tau \in (0,1),$$

observăm că ultimul termen este

$$\frac{\partial f}{\partial x_1}(u + \tau h)h_1 + \dots + \frac{\partial f}{\partial x_m}(u + \tau h)h_m$$

În continuare, notăm $h = u + \tau h$ iar dacă presupunem că h este suficient de mic, atunci rezultă că fiecare derivată parțială $\frac{\partial f}{\partial x_i}(h)$ are același semn ca $\frac{\partial f}{\partial x_i}(u)$. Pe de altă parte, deoarece h_i sunt perturbații pe direcția gradientului, adică

$$h_i = -\eta_i \frac{\partial f}{\partial x_i}(u), \quad \eta_i > 0,$$

rezultă

$$\frac{\partial f}{\partial x_1}(h)h_1 + \dots + \frac{\partial f}{\partial x_m}(h)h_m = -\eta_1 \frac{\partial f}{\partial x_1}(u) \frac{\partial f}{\partial x_1}(h) - \dots - \eta_m \frac{\partial f}{\partial x_m}(u) \frac{\partial f}{\partial x_m}(h) \leq 0$$

astfel încât, obținem:

$$f(u + h) \leq f(u), \quad h = -f'(u)\eta, \quad \eta = (\eta_1, \dots, \eta_m)$$

În concluzie, achizițiile genetice de forma, $u + h$, care sunt realizate pe direcția gradientului, corespund unei performanțe crescute în raport cu minimizarea funcției obiectiv.

Evident, nu putem spune ce înseamnă “ h suficient de mic”, astfel încât raționamentul este mai mult calitativ decât cantitativ. În aceasta constă și marea diferență dintre un algoritm genetic și un algoritm “clasic” de tip Newton. În locul unei alegeri destul de dificile, chiar imposibile, a unei deplasări exacte pe direcția gradientului, un algoritm genetic va elimina deplasările care nu măresc performanța, în baza unui criteriu inspirat din biologie.

Urmașii

Introducerea unor noi indivizi într-o populație se face ținând cont de codificarea folosită. În orice caz, împrumutând mecanismul “cross-over” de combinare a lanțurilor genetice, formal, fiecare individ nou introdus va avea o zestre genetică obținută de la doi “părinți” prin ruperea și refacerea lanțurilor de cromozomi. În cazul problemei noastre, considerăm doi indivizi $u = (x', y')$ și $v = (x'', y'')$, în care $x', x'' \in \mathbb{R}^p$, $y', y'' \in \mathbb{R}^q$, $p + q = m$ iar prin urmași vom considera indivizii (x', y'') și (x'', y') . Anumiți “cromozomi buni” vor fi conservați, și anume aceia care corespund direcției gradientului preluată de la “părinți”, după cum sunt conservați și cromozomii “răi”. În final, prin selecție “naturală” vor rămâne în populație numai cei mai buni indivizi.

Algoritm

Un algoritm evolutiv utilizează noțiunile de populație, generație, selecție, la fel ca în mediul biologic. Un astfel de algoritm poate fi descris prin blocuri, astfel:

Pasul 1. Generează populația inițială $\Pi(t)$, corespunzătoare unui moment inițial t .

Pasul 2. Adaugă noua generație; modifică $t \leftarrow t + 1$;

Pasul 3. Aplică mutații genetice, unor indivizi aleși aleator, pe direcția gradientului.

Pasul 4. Evaluează indivizii prin intermediul funcției obiectiv, f .

Pasul 5. Elimină suprapopulația.

Pasul 6. Continuă cu **pasul 2**, până când este îndeplinită condiția de terminare.

Condiția de încheiere poate să fie reprezentată de un număr maxim de generații sau de o condiție de tip aposteriori de forma:

$$f(u) - f(u+h) < \varepsilon,$$

unde ε este suficient de mic, având semnificația unei precizii date. Ca o paralelă cu mediul biologic, putem să observăm că o condiție de tip aposteriori reflectă, într-un fel, “perfectiunea” adică momentul în care indivizii nu mai pot progresa, în raport cu funcția obiectiv.

Convergeța

Vom presupune că f este mărginită inferior, iar derivata sa satisface o condiție de tip Lipschitz:

$$\|f'(a) - f'(b)\| \leq L\|a - b\|, \quad a, b \in D, \quad L > 0$$

Notăm cu u_n individul cel mai bine plasat în generația “ n ” și rezultă, evident, că $f(u_{n+1}) \leq f(u_n)$ deci șirul $f(u_n)$ este convergent.

Utilizând condiția aposteriori, rezultă:

$$f(u_n) - f(u_n + h) < \varepsilon$$

iar cum \square este suficient de mic, putem considera că $\lim_{n \rightarrow \infty} f(u_n + h_n) = \lim_{n \rightarrow \infty} f(u_n)$.

Pe de altă parte, pentru $u \in D$ și mutațiile genetice $h_i = \eta_i \frac{\partial f}{\partial x_i}(u)$, avem:

$$\begin{aligned} f(u+h) - f(u) &= f'(u + \tau h)h = f'(u)h + (f'(u + \tau h) - f'(u))h \leq \\ & f'(u)h + \|(f'(u + \tau h) - f'(u))h\| \leq f'(u)h + \|f'(u + \tau h) - f'(u)\| \|h\| \leq \\ & - \sum_{i=1}^m \eta_i \left(\frac{\partial f}{\partial x_i}(u) \right)^2 + L\tau \sum_{i=1}^m \eta_i^2 \left(\frac{\partial f}{\partial x_i}(u) \right)^2 = \sum_{i=1}^m \eta_i (-1 + L\tau \eta_i) \left(\frac{\partial f}{\partial x_i}(u) \right)^2 \end{aligned}$$

Pentru η_i suficient de mici, există $\delta > 0$ astfel încât:

$$f(u+h) - f(u) \leq -\delta \|f'(u)\|^2, \quad \text{sau} \quad f(u) - f(u+h) \geq \delta \|f'(u)\|^2$$

ceea ce, împreună cu condiția aposteriori, implică:

$$\delta \|f'(u)\|^2 \leq \varepsilon \tag{1}$$

În sfârșit, pentru $u = u_n$, rezultă $f'(u_n) \rightarrow 0$, $n \rightarrow \infty$.

În continuare, presupunem că f posedă diferențială de ordin 2, și că aceasta este pozitiv definită, adică există $m > 0$ astfel încât:

$$\langle f''(x)y, y \rangle \geq m\|y\|^2$$

Rezultă că există un unic punct de minim al lui f , u_{\min} și utilizând formula lui Taylor există \bar{u} , astfel încât:

$$f(u_{\min}) - f(u_n) - f'(u_n)(u_{\min} - u_n) = \frac{1}{2} \langle f''(\bar{u})(u_{\min} - u_n), (u_{\min} - u_n) \rangle$$

$$f(u_n) - f(u_{\min}) - f'(u_{\min})(u_n - u_{\min}) = \frac{1}{2} \langle f''(\theta)(u_n - u_{\min}), (u_n - u_{\min}) \rangle$$

în care, prin $\langle \cdot, \cdot \rangle$ am notat produsul scalar în P^m .

Folosind ipoteza de pozitiv definire a diferențialei de ordin 2, prin sumarea celor două formule anterioare, rezultă:

$$-f'(u_n)(u_{\min} - u_n) - f'(u_{\min})(u_n - u_{\min}) \geq m \|u_{\min} - u_n\|^2 \quad (2)$$

Deoarece $f'(u_n) \rightarrow 0$ și $f'(u_{\min}) = 0$, obținem $u_n \rightarrow u_{\min}$, $n \rightarrow \infty$

Observăm că din relația (2) rezultă:

$$m \|u_{\min} - u_n\|^2 \leq \|f'(u_n)\| \|u_{\min} - u_n\|, \text{ adică } \|u_{\min} - u_n\| \leq \frac{1}{m} \|f'(u_n)\|$$

În sfârșit, utilizăm relația (1) astfel încât obținem:

$$\|u_{\min} - u_n\| \leq \frac{1}{m} \sqrt{\frac{\varepsilon}{\delta}}$$

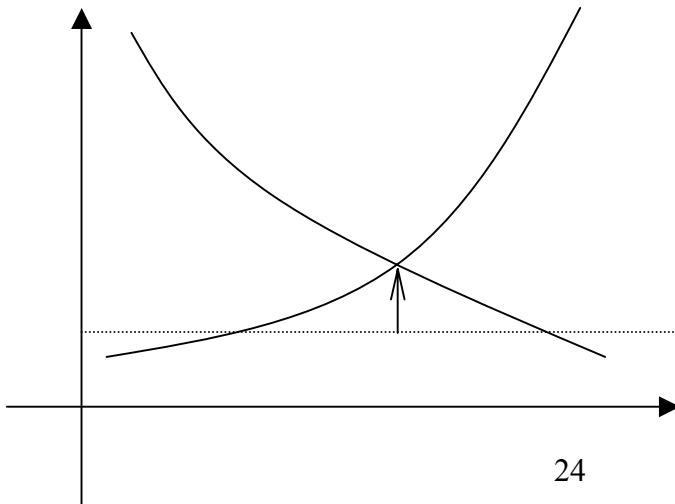
ceea ce arată că utilizarea condiției a posteriori este un criteriu suficient de convergență a șirului (u_n) .

Concluzie

Algoritmul de evoluție converge la valoarea optimă. Există o strânsă legătură între un astfel de algoritm și o metodă clasică de tip *Newton*, ambele utilizând direcția gradientului. Pe de altă parte, există și o mare deosebire între cele două metode concretizată prin “mărimea” deplasării. În metodele de tip *Newton* aceasta este calculată utilizând diferențiala de ordin 2, iar condițiile de aplicare sunt mai mult teoretice decât practice, în vreme ce printr-un algoritm genetic, deplasările “eronate” sunt pur și simplu eliminate, printr-un mecanism de selecție.

Utilizarea unui algoritm genetic pentru determinarea soluției problemei Cauchy, pentru ecuații diferențiale

Există probleme de mare interes pentru cercetarea econometrică, prin care suntem conduși la exprimări care folosesc ecuații diferențiale. De exemplu, echilibrul dintre cerere și ofertă permite o exprimare de acest fel. Să presupunem, ca în figura de mai jos, ca în raport cu cererea și oferta la un moment dat, cererea depășește oferta. În acest caz ne putem aștepta la o creștere a prețului, de exemplu, proporțională cu diferența dintre cerere și ofertă:



$$\frac{\Delta P}{\Delta t} = a(D(t) - S(t))$$

în care a este o constantă pozitivă. Presupunând variații mici ale parametrului t , de exemplu printr-un proces de trecere la limită, rezultă:

$$P'(t) = a(D(t) - S(t))$$

Atunci când cererea depășește oferta, membrul drept al egalității de mai sus este pozitiv, de unde rezultă că funcția P este crescătoare.

În concluzie, un model simplu al variației prețului în funcție de cerere și ofertă, este reprezentat de o problemă *Cauchy*:

$$y' = f(x, y), y(x_0) = y_0,$$

în care x este variabila dependentă, iar $y = y(x)$ este variabila dependentă (endogenă).

Desigur, pentru ca problema să aibă soluție, sunt necesare ipoteze suplimentare asupra problemei. Astfel, dacă funcția:

$$f : [x_0 - X, x_0 + X] \times [y_0 - Y, y_0 + Y] \rightarrow \mathbb{R}$$

este continuă și dacă satisface o condiție de tip *Lipschitz* în raport cu cel de-al doilea argument

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|$$

atunci soluția problemei *Cauchy* există și este unică.

Există mai multe metode de determinare a soluției, unele analitice, dar, în practică se utilizează metode de calcul numeric (aproximativ) cea mai cunoscută fiind metoda *Runge-Kutta*. În această secțiune, vom arăta că determinarea aproximativă a soluției se poate face prin utilizarea unui algoritm de tip genetic.

Adaptarea algoritmului genetic

Soluții admisibile. La fel ca și în cazul metodei *Runge-Kutta*, prin utilizarea unui algoritm genetic, vom determina valorile soluției y în raport cu o diviziune echidistantă a domeniului de definiție al funcției $y = y(x), y : [a, b] \rightarrow \mathbb{R}, x_0 = a < x_1 < \dots < x_n = b, x_i = a + ih, h = \frac{b-a}{n}$.

Vom nota valorile funcției, în punctele diviziunii, prin $y_i = y(x_i), i = 1 \dots n$ astfel încât vectorul (y_1, y_2, \dots, y_n) devine o soluție admisibilă.

Populația. Având ca model populația biologică, vie, vom considera că aceasta este o submulțime a tuturor combinațiilor cromozomiale, adică o submulțime a soluțiilor admisibile. Pentru un moment fixat, notat t , $P(t)$ va reprezenta populația, iar un individ $y = (y_1, y_2, \dots, y_n)$ va fi caracterizat de zestrea sa cromozomială conținută în valorile y_i .

Selecția. La fel ca și în cazul populațiilor biologice, indivizii problemei noastre, pot avea niveluri diferite de adaptare. Această adaptare va fi măsurată prin intermediul unei funcții de performanță. La fiecare moment, surplusul populației va fi eliminat astfel încât populația totală să fie constantă.

Vom considera o formă discretizată a problemei, utilizând aproximarea discretei date de formula:

$$y'(x_i) \approx \frac{y_i - y_{i-1}}{h}$$

$$\left| y'(x_i) - \frac{y_i - y_{i-1}}{h} \right| \leq \text{const} \cdot h$$

Problema *Cauchy* devine:

$$\frac{y_i - y_{i-1}}{h} = f(x_i, y_i), i = 1 \dots n$$

Evident, relațiile anterioare reprezintă un sistem de ecuații neliniare, a cărui soluție (y_1, y_2, \dots, y_n) urmează să o determinăm.

Pentru fiecare soluție admisibilă vom considera formula de eroare:

$$\left(\frac{y_i - y_{i-1}}{h} - f(x_i, y_i) \right)^2$$

iar funcția de performanță asociată vectorului $y = (y_1, y_2, \dots, y_n)$ va fi definită de relația:

$$F(y) = \sum_{i=1}^n \left(\frac{y_i - y_{i-1}}{h} - f(x_i, y_i) \right)^2$$

Vom considera că un individ din populația $P(t)$ este cu atât mai bien adaptat cu cât funcția de performanță are o valoare mai mică.

Urmașii. Doi indivizi pot da naștere la alți doi indivizi, prin combinarea zestrei lor cromozomiale. Astfel, printr-o operație *cross-over*, indivizii (y_1, y_2, \dots, y_n) și (z_1, z_2, \dots, z_n) vor da naștere urmașilor:

$$(y_1, y_2, \dots, y_k, z_{k+1}, \dots, z_n) \text{ și } (z_1, z_2, \dots, z_k, y_{k+1}, \dots, y_n)$$

Mutații genetice. Orice cromozom poate să sufere mutații, întâmplătoare sau, dacă admitem că educația are un rol, putem accepta și anumite achiziții genetice. Indiferent de situația, vom modela o mutație genetică a cromozomului y_i sub forma $y_i \pm \varepsilon$.

Convergența algoritmului

Notăm cu u_t individul din populația $P(t)$ având cea mai mică valoare a funcției de eroare F . Așa cum am demonstrat în secțiunea anterioară, șirul $(u_t)_{t \geq 0}$ este convergent, iar limita sa este soluția problemei de optimizare $\inf F$.

Utilizând convergența $u_t \rightarrow u$, rezultă că pentru orice $\varepsilon > 0$, există $y = (y_1, y_2, \dots, y_n)$ astfel încât:

$$F(y) = \sum_{i=1}^n \left(\frac{y_i - y_{i-1}}{h} - f(x_i, y_i) \right)^2 < \varepsilon$$

În consecință, pentru $h > 0$, printr-o eventuală renotare a valorilor, rezultă că există $y = (y_1, y_2, \dots, y_n)$ astfel încât:

$$\left| \frac{y_i - y_{i-1}}{h} - f(x_i, y_i) \right| < h$$

Având în vedere formula de aproximare a derivatei, rezultă:

$$|y'(x_i) - f(x_i, y_i)| \leq \left| y'(x_i) - \frac{y_i - y_{i-1}}{h} \right| + \left| \frac{y_i - y_{i-1}}{h} - f(x_i, y_i) \right| < Ch$$

ceea ce arată că valoarea finală a algoritmului, și anume $u_t = y = (y_1, y_2, \dots, y_n)$ reprezintă o aproximare a soluției problemei *Cauchy problem*, cu precizia h .

Algoritmul

Prezentăm în continuare adaptarea algoritmului general din prima secțiune, pentru determinarea aproximativă a soluției problemei *Cauchy*.

Pasul 1. Generează populația inițială, sub forma unei mulțimi de M vectori, având ca valori, perturbații uniforme ale valorii inițiale a problemei $(y_0 \pm \varepsilon_1^m, y_0 \pm \varepsilon_2^m, \dots, y_0 \pm \varepsilon_n^m)$, $m = 1..M$, $\varepsilon_i^j > 0$.

Pasul 2. Generează urmașii și incrementează contorul de timp, $t \leftarrow t + 1$. Presupunem că numai indivizii cei mai bine adaptați participă, iar pentru (y_1, y_2, \dots, y_n) și (z_1, z_2, \dots, z_n) selectăm cromozomii având cele mai apropiate valori, respectiv indicele k pentru care $|y_k - z_k| = \min_{i=1..n} |y_i - z_i|$.

Printr-o operație *cross-over*, avem $(y_1, y_2, \dots, y_k, z_{k+1}, \dots, z_n)$ și $(z_1, z_2, \dots, z_k, y_{k+1}, \dots, y_n)$.

Pasul 3. Aplică mutații genetice pentru o noua generație, prin perturbații pe direcția gradientului funcției de eroare $F(y_1, y_2, \dots, y_n) = \sum_{i=1}^n \left(\frac{y_i - y_{i-1}}{h} - f(x_i, y_i) \right)^2$. După cum am demonstrat în prima parte, astfel de mutații sunt esențiale pentru convergența algoritmului.

Pasul 4. Clasifică indivizii, în funcție de performanța măsurată prin intermediul funcției de eroare F .

Pasul 5. Elimină surplusul de populați, astfel încât aceasta să păstreze un număr constant de indivizi.

Pasul 6. Continuă cu **pasul 2**, pînă la îndeplinirea condiției de oprire, de tip a posteriori:

$$F(u_t) - F(u_{t-1}) < h,$$

în care u_t reprezintă individul cu cea mai mare performanță din populația $P(t)$.

Concluzie.

Determinarea soluției unei probleme *Cauchy* se poate face printr-un algoritm genetic, valorile determinate fiind relative la o diviziune, la fel ca și în cazul metodei *Runge Kutta*.

Am aplicat algoritmul descris pentru problema *Cauchy* $y'(t) = 2(t^3 - t^2)$, $y(1) = 0$, $y: [1, 2] \rightarrow \mathbb{R}$. Populația considerată are 30 de indivizi, clasificați în funcție de performanță, iar la fiecare pas am ales primii 20 de indivizi, care generează 20 de urmași. Pentru $h = \frac{1}{10}$, am obținut soluția y , cu două zecimale exacte:

$$(0, 0.01, 0.05, 0.13, 0.26, 0.45, 0.71, 1.07, 1.53, 2.11, 2.83)$$

Bibliografie

1. Albu L.L, Macroeconomie non-lineara si prognoza, Academia Romana, 2002
2. Banzhaf W., Nordon P., Keller R.E., Francone F.D., Genetic Programming – An introduction, Morgan Kaufmann Publishers, San Francisco 1998
3. Mateescu G.D., Optimization by using evolutionary algorithms with genetic acquisitions, Romanian Journal of Economic Forecasting, 2/2005
4. Stoer J., Bulirsch R., Introduction to Numerical Analysis, Springer-Verlag, 1992