

ODHADY PARAMETRŮ EKONOMICKÝCH MODELŮ POMOCÍ GENETICKÝCH ALGORITMŮ

M. Ševela

Došlo: 15. prosince 2003

Abstract

ŠEVELA, M.: *Applicability of genetic algorithms to parameter estimation of economic models*. Acta univ. agric. et silvic. Mendel. Brun., 2004, LII, No. 3, pp. 79-86

The paper concentrates on capability of genetic algorithms for parameter estimation of non-linear economic models. In the paper we test the ability of genetic algorithms to estimate of parameters of demand function for durable goods and simultaneously search for parameters of genetic algorithm that lead to maximum effectiveness of the computation algorithm. The genetic algorithms connect deterministic iterative computation methods with stochastic methods. In the genetic algorithm approach each possible solution is represented by one individual, those life and lifes of all generations of individuals run under a few parameter of genetic algorithm. Our simulations resulted in optimal mutation rate of 15% of all bits in chromosomes, optimal elitism rate 20%. We can not set the optimal extend of generation, because it proves positive correlation with effectiveness of genetic algorithm in all range under research, but its impact is decreasing. The used genetic algorithm was sensitive to mutation rate at most, than to extend of generation. The sensitivity to elitism rate is not so strong.

genetic algorithm, non-linear models, iterative algorithm, demand function

Ekonomická teorie a praxe již dlouhou dobu nevy-
stačí s analýzou pouze lineárních vztahů. Neustále
rostoucí komplexnosti modelů odpovídá i rostoucí
složitost vzájemných vztahů. V mnoha případech
již není možné spoléhat pouze na lineární závislosti
proměnných či na dostatečnou přesnost aproximace
nelineárního vztahu vztahem lineárním. Modely
se vztahy, které jsou lineární v odhadovaných pa-
rametrech, byly předmětem intenzivního výzkumu od
30. let minulého století, jejich analýza je základním
stavebním kamenem současné ekonometrie. Odhady
hledaných parametrů lineárních modelů lze zpravidla
získat řešením soustav lineárních rovnic vyplývajících
z podmínek extrému zvolené kritériální funkce.
Při splnění vybraných statistických předpokladů jsou
tyto odhady nejlepšími nestrannými odhady. Výše
uvedených výhod lze využít i u vztahů původně neli-

neárních, které lze vhodnou transformací (např. loga-
ritmováním) převést na vztahy lineární v parametrech
(Hušek, 1989, Stundenmund, 2002).

Odhady parametrů modelů nelineárních v paramet-
rech vychází z obdobným principů jako modely line-
ární, podmínky dosažení extrému kritériální funkce
jsou ale popsány soustavou nelineárních rovnic.
Řešení této soustavy není triviální a často nelze řešení
jednoznačně vypočítat. Často nelze ani nalézt analy-
tické vyjádření podmínek dosažení extrému kritéri-
ální funkce, neboť neumíme vyjádřit příslušné první
parciální derivace. V těchto případech musíme prac-
ovat přímo s kritériální funkcí a hledat její extrém. Hle-
daným extrémem kritériální funkce bývá často mini-
mum, zejména při využití kritéria nejmenších čtverců
či preidentifikované metody zobecněných momentů.
Hledaným extrémem může být i maximum, pokud

použijeme odhadové kritérium maximální věrohodnosti (Verbeerk, 2002, Perasan, 2000).

Odhady parametrů se získávají různými numerickými metodami, obvykle iteračního charakteru. Tyto metody se vyznačují velkou výpočetní náročností, jejich aplikace je ve větším rozsahu podmíněna nasazením současných výpočetních technik. I přes podstatné zrychlení výpočtů nelze procházet celý několikarozměrný definiční obor kritériální funkce. Stále mají svůj význam optimalizační algoritmy a jiné odhadové postupy, které značně snižují výpočetní náročnost iteračních metod.

Většina používaných iteračních metod pracuje na gradientním principu. Gradientní metody deterministicky postupují od zvoleného počátečního řešení směrem k optimu. Jejich postup skrývá nebezpečí nalezení pouze suboptimálního řešení, které je pouze lokálním extrémem. Pokud kritériální funkce není hladkou, zůstanou tyto metody v prvním nalezeném údolí, v němž sestoupí zcela dolů. Současně mohou některá údolí překročit, pokud je délka kroku příliš velká. Vzniká zde velké riziko nalezení nesprávného bodu extrému, které je způsobeno systematicky chybnou volbou počátečního řešení. Navíc zmíněné iterační algoritmy nedokážou chyby tohoto druhu žádným způsobem indikovat (Rektorys, 1964, Hušek, 1999, Limpouch, 2003).

Jedním z řešení je použití stochastických výpočetních postupů, které dokážou uvedené chyby zmírnit využitím principu náhodných výběrů a opakovaných výpočtů. Stochastické výpočetní postupy jsou výpočetně více náročné než deterministické, ale poskytují obvykle spolehlivější výsledky. I při stochastických výpočtech je nutno použít optimalizované postupy, které celý proces značně urychlí. Stochastické algoritmy jsou složitější než deterministické, jsou řízeny podstatně větším množstvím různých parametrů. Správná volba parametrů podstatně ovlivňuje nejen spolehlivost poskytovaných výsledků, ale i rychlost konvergence odhadů.

Optimální volba výpočetních parametrů stochastických algoritmů není zcela jednoznačná a silně závisí nejen na samotném algoritmu, ale přímo i na funkci, u níž hledáme globální extrém. Parametry nelze určit zcela jednoznačně, jejich volba v určitém intervalu poskytuje přibližně stejné výstupy stochastického algoritmu, měřeno nejen rychlostí konvergence odhadů, robustností odhadů či výpočetní náročností. Obvykle jsou tyto sledované vlastnosti navíc substituty. Používané procesy stochastických algoritmů jsou často sestaveny na základě podobného přírodního mechanismu. Nejinak tomu není u genetických výpočetních postupů, které budou použity v této práci.

Cílem práce je ověřit vhodnost genetických algoritmů pro hledání extrému zmíněných funkcí a

odvodit optimální výpočetní parametry nebo alespoň vhodné intervaly hodnot parametrů. Při odvozování parametrů modelů chování spotřebitelů a dalších ekonomických subjektů je obvykle použito kritéria nejmenších čtverců nebo kritéria maximální věrohodnosti při předpokladu normálního rozdělení pravděpodobnosti. Kritériální funkce nejmenších čtverců bude předmětem aplikace genetických algoritmů na problematiku odvození parametru funkce poptávky po statcích dlouhodobé spotřeby.

Principy genetických algoritmů

Genetické algoritmy jsou jednou z oblastí evolučního programování, které se vyvíjí jako samostatná disciplína v rámci přístupů umělé inteligence. Počátky evolučního programování se datují do 60. let 20. století, kdy I. Rechenberg publikoval první výsledky, v dnes již klasickém díle evolučního programování, *Evolutionsstrategie* (Rechenberg, 1965). První genetické algoritmy byly sestaveny J. Hollandem a publikovány v roce 1975. Od těchto počátků prošly genetické algoritmy dlouhým vývojem a řadou úspěšných aplikací. Okruh problémů, pro které je použití genetických algoritmů přínosné, je neustále rozšiřován. Na konci minulého století probíhaly dokonce pokusy využití postupů genetických algoritmů pro tvorbu počítačových programů v hierarchickém jazyce LISP (Koza, 1992).

Genetické algoritmy pracují na principu genetických zákonů a dalších evolučních zákonitostí. Využívají velice jednoduchého evolučního principu přírody, kdy ze skupiny jednotlivců přežívají ti lepší na úkor slabších. Jedinci vykazující lepší vlastnosti se stávají ve skupině dominantní, mají více potomků a posouvají tak celou skupinu směrem k lepšímu přizpůsobení. Všechny tyto procesy přirozeného výběru a stále větší adaptace samozřejmě probíhají pod vlivem nahodilých událostí, které mohou přinášet nové důležité prvky do vývoje celé skupiny.

Genetické algoritmy lze používat v roli vyhledávacích algoritmů, jejichž úkolem kterých je nalézt nejlepší řešení z daného souboru. Postupným evolučním vývojem je hledáno řešení, které vyhovuje zadaným podmínkám lépe než řešení ostatní. Tento přístup k hledání nejlepšího řešení se odlišuje od obvyklých metod zejména v následujících bodech (Sean, 2003, Tierney, 2003).

- Není vyšetřováno pouze jedno v úvahu připadající řešení, ale je zkoumáno více možných řešení současně (tj. celá populace jednotlivých řešení).
- Není třeba znát vlastnosti zkoumaného kauzálního vztahu mezi vstupy a výstupem, umět identifikovat podezřelé body atd. Zcela postačuje schopnost jednoznačného porovnání dvou řešení v ordinalistickém smyslu (výstup nemusí být měřitelný).
- Nepracuje se přímo s řešeními, ale jejich kódy.

Tento způsob reprezentace vstupních dat je velice vhodný například pro analýzu pořadí bodů při řešení problému obchodního cestujícího, popř. pro analýzy diskrétních vstupních veličin.

- Postup řešení není deterministický, nejsou posuzována všechna možná řešení. Při výběru posuzovaných řešení se využívá záměrně zkreslených pravděpodobnostních přístupů.

Výsledky genetických výpočetních algoritmů jsou podobně jako u ostatních iteračních postupů ovlivněny zvoleným prvotním řešením, které je následně optimalizováno. Při používání genetických výpočetních postupů je ale riziko nevhodně zvoleného počátečního řešení podstatně sníženo náhodnou volbou celé generace řešení. Dále simulací procesu mutace dochází k náhodnému vzniku zcela nových jedinců, kteří mohou přinést do řešení nové, dosud neprozkoumané vlastnosti a stát se tak základem vhodné genové kombinace. Tyto mechanismy vedou k prozkoumání daleko větší oblasti možných řešení než postupy obvyklé a snižují tak riziko uvážnutí v suboptimálním řešení.

Celý výpočetní postup lze popsat sekvencí po sobě následujících kroků, které vyjma kroku číslo 1 probíhají opakovaně buď do předem zadaného počtu simulovaných generací nebo do počtu generací bez zlepšení nalezeného řešení.

1. **Vytvoření první generace jedinců** zcela náhodným výběrem z definičního oboru zkoumaného vztahu. Genetické algoritmy pracují obvykle s konstantním počtem jedinců v jedné generaci. Příliš vysoké rozsahy jedné generace nevedou sice k vyšší efektivnosti výpočtu, ale nejsou dostatečně efektivním substitutem k podstatně delším dobám výpočtu. Neúměrně malé počty jedinců způsobují ztrátu diverzity a pomalou konvergenci k optimálnímu řešení (Obitko, 1998).

2. **Ohodnocení kvalit jednotlivých řešení** často spojené se seřazením jednotlivých řešení podle tohoto kritéria. Jak již bylo zmíněno, na hodnotící funkci je kladem pouze požadavek souladu s axiomem úplnosti srovnání, což umožňuje řešení širokého spektra problémů.

3. **Výběr jedinců – rodičů – pro křížení** prostřednictvím zkresleného náhodného výběru. Jednotlivým řešením ve stávající generaci je přiřazena pravděpodobnost výběru podle jejich kvality. Existuje několik základních principů, které využívají úměry k pořadí nebo v případech měřitelnosti kvality jedinců i úměry k dosud nejlepšímu nalezenému řešení.

4. **Tvorba genotypu nového jedince křížením** genové výbavy rodičů. Nový jedinec získává část genové výbavy rodičů v každém svém chromozomu. Pokud jedinec nespadá do definičního oboru zkoumaného problému, hledá se

křížením stejných rodičů další jedince. Takto potomek získá náhodnou část genové výbavy rodičů, která se ale může projevit zcela odlišným chováním nového jedince.

5. **Mutace genotypu** představuje mechanismus simulace náhodných poruch při křížení, který napomáhá zachování diverzity populace a přináší zcela nová, dosud neprozkoumaná řešení. Příliš velká míra mutací vede k pomalejší konvergenci výsledků k optimálnímu řešení, současně představuje snížení rizika neprozkoumání částí definičního oboru zkoumaného problému.

6. **Nahrazení staré generace novou generací** není obvykle úplné. Mohlo by tak dojít ke ztrátě nejlepších řešení a nejlepšího genetického materiálu. Míra zachování původní generace – elitismu – podstatně zlepšuje výkonnost genetických algoritmů. Na její výši závisí proměnlivost jednotlivých generací a tím i maximální rychlost adaptace celé generace jedinců.

Proces a výsledky genetických algoritmů jsou ovlivňovány několika základními parametry. Při hledání nejlepšího řešení dochází ke zkreslenému stochastickému zkoumání definičního oboru. Čím větší část definičního oboru algoritmus prozkoumá, tím můžeme řešení považovat za spolehlivější. Je patrné, že kvalita řešení je přímo úměrná počtu simulovaných generací, velikosti generace a míře mutací. Vliv míry elitismu není jednoznačný, neboť umožňuje zachovat stávající nejlepší výsledky, ale na druhou stranu snižuje rychlost adaptace nových generací.

Simulace genetických algoritmů jsou i při současných parametrech výpočetních systémů značně časově náročné. Nelze tedy simulovat nekonečně velké počty generací či jedinců pro nalezení vysoce spolehlivého řešení. Při hledání řešení je tedy nutno i nadále optimalizovat samotný proces hledání řešení, aby uspokojivý výsledek byl nalezen v co nejkratším čase. Toho lze dosáhnout zejména zúžením definičního oboru na základě apriorních informací, omezení přesnosti výpočtu zejména u spojitých veličin atd. Další nezanedbatelnou možností je nalezení vhodných parametrů genetického algoritmu, zejména míry mutací a elitismu.

Jako každý výpočet vyžadující počáteční řešení, jsou i výsledky genetických algoritmů závislé na tomto parametru. Přestože je jeho dopad snížen náhodnou volbou celého souboru počátečních řešení, stále se jeho vliv projevuje. Vhodnou možností dalšího snížení jeho vlivu je opakování celé simulace s jiným náhodně vytvořeným souborem počátečních řešení. Pokud se výsledky řešení simulací neliší a konvergují ke stejným hodnotám jak v rámci jednotlivých simulací, tak i simulací při různých počátečních řešení, lze řešení získané genetickým algoritmem považovat za spolehlivé.

Přístupy a metodika měření výsledků odhadů

Hodnocení přesnosti a spolehlivosti výsledků genetických algoritmů není zcela bez problémů. Genetické algoritmy nejsou deterministickými výpočetními postupy, jejich výsledky mají stochastický charakter. Opakované simulace dospívají k rozdílným výsledkům nejen z důvodu náhodně generovaného souboru počátečních řešení, ale i z důvodu aplikace náhodných postupů v procesu křížení chromozomů a procesu mutace. Je tedy vhodné provádět opakované simulace a výsledky zhodnotit použitím statistických metod.

K nepříznivým skutečnostem patří existence rozdílné míry vlivu jednotlivých vstupních proměnných na výslednou hodnotu kritériální funkce. Malá změna v hodnotě významné proměnné může být doprovázena značným odchylením méně vlivných proměnných, neboť představují menší penalizaci při výpočtu hodnoty kritériální funkce. Dojde-li ke změně tohoto typu v posledních generacích simulace, mohou tyto odchylky přetrvat do výsledného řešení. Při procesu simulace tedy nejdříve musí dojít ke stabilizaci významnějších proměnných, pak konvergují i proměnné méně významné. Z tohoto důvodu je obvykle dosahováno větší přesnosti u významnějších proměnných. Tento negativní jev lze zmírnit vhodnou substitucí proměnných ve zkoumané funkci, alespoň na úroveň stejného řádu vysvětlujících proměnných. Dalším možným způsobem snížení tohoto negativního dopadu je využití průměru hodnot za několik posledních generací za předpokladu oboustranně symetrického přibližování k optimální hodnotě.

Dalším nepříznivým jevem je možnost vzájemné kompenzace hodnot vstupních proměnných. Odchylka jedné proměnné je často kompenzována opačnými odchylkami ostatních proměnných. Z tohoto důvodu je vhodné také posuzovat současnou konvergenci všech proměnných. Možnost vzájemné kompenzace vstupních proměnných je dalším argumentem pro provádění opakovaných simulací. Při větším počtu měření dochází k symetrickým odchylkám na obě strany, pomocí ukazatelů polohy lze následně učinit nestranný odhad a konstruovat intervaly spolehlivosti.

Rychlost a přesnost konvergence závisí na velikosti prohledávaného definičního oboru funkce. Z počátku samozřejmě neznáme ani přibližně velikosti hledaných odhadů. Výpočet pomocí genetického algoritmu je vhodné provést dvoustupňovitě. V prvním kroku povolíme maximální technicky přípustné hodnoty a získané odhady poslouží k omezení prohledávaného prostoru při výpočtu ve druhém kroku. Lze samozřejmě testovat hodnoty vytvořených binárních reprezentací chromozomů a jejich tvorbu popř. opakovat. Tento postup je ale neefektivní a může vést k výraznému zpomalení algoritmu při dlouhých pseudoná-

hodně generovaných binárních řetězcích a malých reálných hodnotách omezujících definiční obor. Na druhou stranu lze takto zajistit splnění apriorních požadavků na odhadované parametry (např. nezápornost). Daleko přínosnější je omezit délku řetězce binární reprezentace chromozomu a tím podstatně urychlit veškeré operace křížení a mutace. Pokud bude provedeno nesprávné omezení prohledávaného prostoru, odhad parametru bude konvergovat k jeho hranici.

Pro funkce, u nichž nastává globální extrém ve více místech (např. některé cyklické funkce), nedojde ke konvergenci vysvětlujících proměnných. Tento problém nebude indikován statistikami polohy, ale statistikami variability, zejména variačním koeficientem.

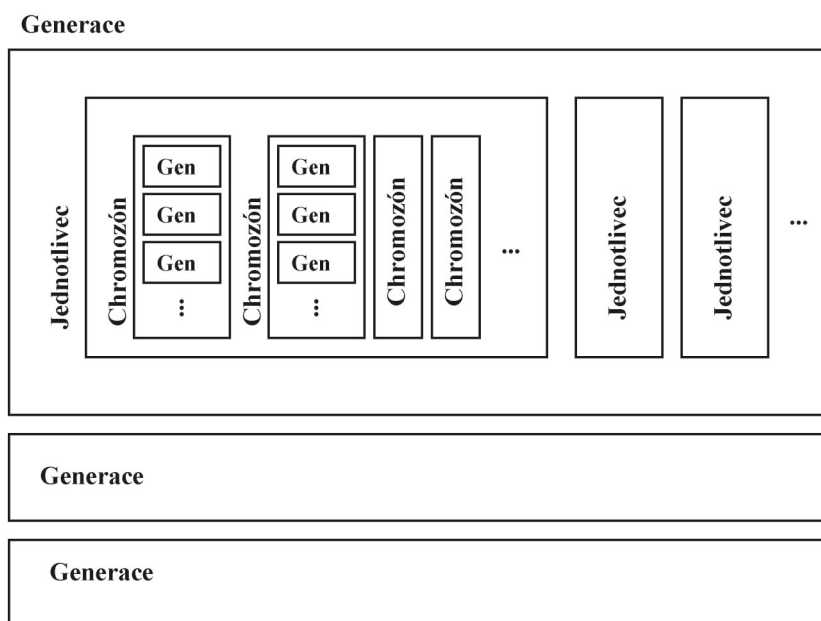
Pro hodnocení kvality a konstrukci odhadů parametrů bude současně použito více rozdílných měř polohy a variability. Mezi používané míry polohy je zařazen aritmetický průměr, který dokáže popsat polohu oboustranně symetricky rozdělených hodnot. Současně ale aritmetický průměr nepatří mezi robustní statistiky, proto bude dále využito i mediánu a aritmetického průměru souboru zmenšeného o horní a dolní decil. Ke všem mírám polohy bude současně vypočten i rozptyl a variační koeficient.

Implementace genetického algoritmu v jazyce Java

Pro implementaci genetických algoritmů za účelem výpočtu globálních extrémů reálných funkcí na omezené množině byl vybrán programovací jazyk Java. Programy v jazyce Java jsou snadno přenositelné mezi jednotlivými platformami a současně tento jazyk poskytuje nástroje pro velice snadnou a bezpečnou práci s objekty. Reprezentace dat používaných při genetické simulaci pomocí objektového přístupu je velmi přímočará. Celou datovou strukturu znázorňuje Obr. 1 – Struktura dat genetické simulace.

Z obrázku je patrná hierarchická struktura objektů. Nejdůležitějším objektem je jednotlivec, který odpovídá jednomu posuzovanému řešení. Mezi základní vlastnosti jednotlivce patří fenotyp reprezentovaný přiřazenou funkční hodnotou a genotyp uložený v jednotlivých chromozomech. Chromozomy jsou rovněž objekty. Reprezentují příslušné vstupní hodnoty kritériální funkce, jejich počet odpovídá počtu vstupních proměnných. Vlastnostmi samotných objektů chromozomů nejsou pouze hodnoty vstupních proměnných kritériální funkce, ale současně i jejich kódované formy, která je obdobou DNA řetězců. Právě na úrovni DNA řetězců probíhá křížení a mutace jednotlivých řešení.

Mechanismus kódování, křížení a mutace lze simulovat různými způsoby. Jejich volba je samozřejmě odvislá od řešeného problému a typu používaných dat. V některých případech je výběr jednoznačný, u jiných

1: *Struktura dat genetické simulace*

Lze aplikovat více možností. Při práci s reálnými čísly se jako optimální osvědčilo binární kódování proměnných, které je současně velice výhodné i z hlediska operací křížení a mutace (Schlutz, 2003). Při tvorbě genetického algoritmu pro výpočet extrémů funkce bylo použita binární reprezentace hodnot. Binárně zapsané hodnoty jsou obvykle celočíselné, požadovaný počet desetinných míst je nahrazen posunem desetinné čárky výsledné desítkové reprezentace binárního řetězce kódu.

Tvorba chromozomů nových jedinců probíhá křížením, kdy je použita náhodná část chromozomu jednoho rodiče a doplněna genetickou informací druhého rodiče tak, aby kód chromozomu byl úplný a správný. V algoritmu bylo použito jednobodové křížení, kdy dochází k vzniku chromozomů nových jednotlivců výměnou genetické informace rodičů od náhodně generované pozice binárního řetězce. Chromozomy prvního potomka jsou tedy totožné s chromozomy prvního rodiče až po náhodně generovanou pozici, zbytek pak tvoří část chromozomu druhého rodiče. U druhého potomka je situace přesně opačná.

Mutace v implementovaném algoritmu probíhá na zcela náhodném základě. Pro každý bit binární reprezentace hodnot v chromozomech je generováno pseudo-náhodné číslo. Pokud toho náhodné číslo překročí vstupní parametr míry mutací, dojde k inverzi daného bitu. Tímto způsobem je simulována možnost poruchy křížení v jakémkoliv bodě procesu. Po ukončení procesu mutace následuje kontrola příslušnosti hodnoty vytvořeného chromozomu do definičního oboru reálné funkce. Pokud je hodnota mimo tento definiční

obor, opakuje se celý proces křížení a mutace daného chromozomu.

Při vytváření nové generace je nejprve zkopírována část nejlepších jedinců ze staré generace, jejíž relativní velikost vzhledem k rozsahu generace je dána mírou elitismu. Pro zbývající volné pozice v generaci jsou vytvářeni potomci. Rodiče jsou vybíráni náhodně z celé staré generace, pravděpodobnost výběrů je nepřímo úměrná pořadí sestavenému podle hodnoty zkoumané reálné funkce. Vždy dva potomci dvou rodičů jsou pak výsledkem výše popsanych procesů křížení a mutace. Počet simulovaných generací odpovídá dalšímu zadávanému parametru genetického algoritmu. Výstupem evolučního procesu sestávajícího ze zadaného počtu generací je soubor elitních jedinců, kteří by byli v nezměněné podobě kopírováni do případné další generace. Tento výstup je v závěru statisticky zpracován za účelem lepšího posouzení dosazených výsledků.

Odhady parametrů poptávkové funkce

Za nelineární formu poptávky byla zvolena poptávková funkce používaná při analýzách poptávky po statcích dlouhodobé spotřeby. Tato poptávková funkce vychází z růstového logistického modelu, kde velikost poptávky za určité časové období je dána rozdílem v relativním vybavení domácností sledovaným statkem. Funkce má následující specifikaci:

$$\frac{dy(t)}{dt} = by(t) - \frac{b}{S} I^2(t)(1 + ct^{-a}),$$

kde $V(t)$ – relativní vybavenost statkem
 S – relativní nasycení při daném příjmu
 a – důchodová elasticita statku
 b – autonomní časový růst
 c – růst příjmů ve sledovaném období.

V takto specifikovaném modelu jsou veličiny a , b , S odhadovanými parametry. Odhady parametrů budou stanovovány použitím kritéria minimální čtvercové odchylky při využití genetického algoritmu.

Cílem je nalézt optimální velikosti parametrů genetického algoritmu. Z tohoto důvodu je nutno stanovit kritérium hodnocení výsledků poskytnutých genetickým algoritmem. Nelze sledovat výslednou chybu odhadu, neboť neznáme skutečné velikosti parametrů, ani ji nejsme schopni odvodit jiným způsobem. Nezbyvá než sledovat míru konvergence odhadů simulací při rozdílných počátečních řešeních a spolehat na minimalizaci chyb odhadů při opakovaných náhodných pokusech.

Za účelem hodnocení výsledků simulací za rozdílných parametrů genetického algoritmu je nutno shrnout výsledky všech simulací při jednom nastavení parametrů do jednoduchého ukazatele. Snadné bude využití variačního koeficientu, který lze snadno vypočítat a interpretovat pro používané ukazatele polohy. Průměr variačního koeficientu přes všechny proměnné a simulace za stejných parametrů poskytne příslušné kritériální hodnoty pro vzájemnou komparaci efektivit genetického algoritmu.

Simulace genetického algoritmu probíhaly pro hodnoty míry mutace a elitismu vždy v rozsahu 0,05 až 0,50 s krokem 0,05. Spodní hranice obou ukazatelů musí být nenulová, aby algoritmus mohl využít předností genetických algoritmů. Horní hranice zkou-

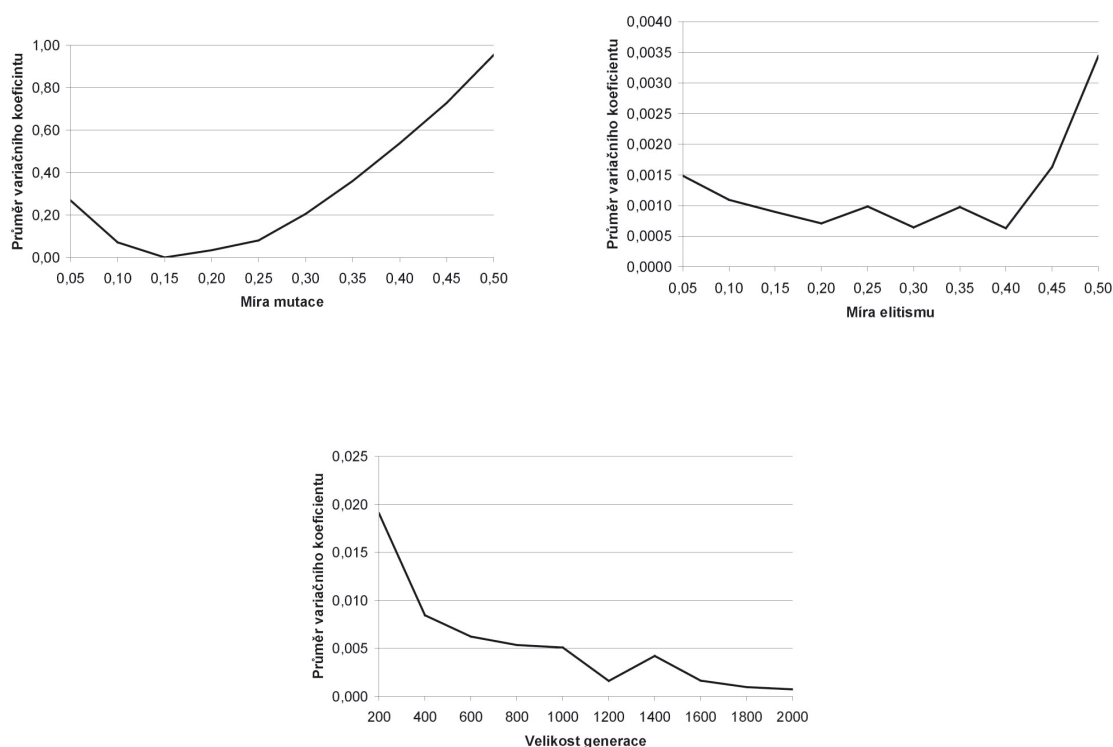
maných intervalů jsou empiricky zjištěnými hodnotami, kdy genetický algoritmus opět přechází v jiný výpočetní postup (Schultz, 2003). Počet jednotlivců v generaci byl simulován v rozsahu 200 až 2000 s krokem 200 jedinců. Spodní a horní hranice byla opět stanovena na základě empirických hodnot (Sean, 2003). Simulace probíhaly při 100 různých počátečních úplných generací řešení. Délka každé simulace byla 100 generací.

Výsledky simulací za jednotlivých parametrů genetického algoritmu nelze prezentovat souborně, neboť se jedná o čtyřrozměrné vektory. Ze setřídění dosažených výsledků podle průměru variačního koeficientu zcela jednoznačně vyplývá, že nejprínosnější jsou hodnoty míry mutace o velikosti 0,15 (u nejlepších tří set výsledků se tato hodnota vyskytuje 98 krát), následovány velikostí 0,10 (79 výskytů v nejlepších měření). U míry elitismu nejsou výsledky takto jednoznačné. Největší zastoupení u tří set nejlepších měření má hodnota 0,20 (38 výskytů), hodnota 0,25 (33 výskytů) a hodnota 0,10 (32 výskytů). Velikost generace jednotlivců vykazuje pozitivní korelaci s kvalitou simulace po celou dobu svého růstu. Příspěvek ke kvalitě ale stále klesá. Optimální velikost generace nelze tady jednoznačně stanovit. Je nutno uvážit přínos velkých generací v komparaci k rostoucím alternativním nákladům ve formě výpočetní náročnosti, která je s velikostí generace v přímé souvislosti.

Vlivu změn v jednotlivých parametrech při podmínce *ceteris paribus* jsou znázorněny v následujících grafech. Hodnoty ostatních parametrů jsou na nejlepší nalezené úrovni (míra mutace – 0,15; míra elitismu 0,20; velikost generace 2000).

I: Absolutní četnosti řešení u souboru nejlepších tří set řešení podle jednotlivých parametrů

Míra mutace	Absolutní četnost	Míra elitismu	Absolutní četnost	Velikost generace	Absolutní četnost
0,05	65	0,05	31	200	10
0,10	79	0,10	32	400	19
0,15	98	0,15	31	600	28
0,20	34	0,20	38	800	31
0,25	24	0,25	33	1000	30
		0,30	30	1200	33
		0,35	30	1400	35
		0,40	28	1600	27
		0,45	26	1800	38
		0,50	21	2000	39



2–4: Citlivost variačního koeficientu na parametry genetického algoritmu

Z výše uvedených grafů je patrné, že při optimální velikosti parametrů je zkoumaný genetický algoritmus nejvíce citlivý na změnu míry mutací, následně na velikost generace. Citlivost na změny míry elitismu není v okolí optima výrazná.

Nalezené nejlepší hodnoty se příliš neliší od hodnot uváděných jako doporučené v pracích Obítka (2003) či Schultz (2003), spadají rovněž do intervalů doporučených pro obecnější úlohy genetického programování (Sean, 2003, Tierney, 2003).

Závěr

Genetické algoritmy jsou jednou z numerických metod, které rozvojem výpočetní techniky nalézají uplatnění v ekonomické analýze. Genetické algoritmy umožňují získávat řešení problémů, která nejsme

schopni nalézt pomocí standardních matematických postupů. Spojení deterministických postupů se stochastickými dokáže poskytovat i spolehlivější řešení než samotné deterministické iterační postupy.

Cílem příspěvku bylo ověřit možnost aplikace genetických algoritmů pro výpočet regresních parametrů nelineárních funkcí. Provedené výpočty prokázaly vhodnost genetického algoritmu, použitý postup poskytoval statisticky vydatné odhady parametrů poptávkové funkce po statcích dlouhodobé spotřeby. Současně byly zjišťovány optimální parametry genetického algoritmu pro zkoumaný typ regresní funkce. Nalezené optimální hodnoty výpočetních parametrů leží přibližně ve středu v intervalů doporučených hodnot platných pro obecnější tvary kritériálních funkcí.

SOUHRN

Příspěvek se zabývá možnostmi aplikace genetických algoritmů na odhady parametrů nelineárních modelů. Cílem příspěvku je ověřit možnost aplikace genetických algoritmů na odhady parametrů poptávkové funkce po předmětech dlouhodobé spotřeby a současně nalézt optimální velikost parametrů genetického algoritmu a tím docílit jeho maximální efektivity. Genetické algoritmy spojují deterministické iterační postupy se stochastickými metodami. Zkoumaná řešení problému jsou ztotožněna s jednotlivci, jejichž život a život celých generací probíhá pod vlivem několika základních parametrů. Z provedených měření vyplynulo, že optimální mírou mutací je 15 % bitů jednotlivých genů, optimální velikostí elitismu je 20 % jednotlivců v každé generaci. Optimální velikost generace nelze stanovit, neboť rostoucí velikost generace vykazuje stále pozitivní, i když klesající vliv na kvalitu odhadů. Při optimální velikosti parametrů je zkoumaný genetický algoritmus nejvíce citlivý na změnu míry mutací, následně na velikost generace. Citlivost na změny míry elitismu není v okolí optima výrazná.

genetický algoritmus, nelineární modely, iterační metody, funkce poptávky

LITERATURA

- ANONYM: Build a genetic algorithm for MaxOne problem. [on-line]. [cit. 12.12.2003]. Dostupné na <http://www.cs.umd.edu/projects/plus/ec/ecj/docs/tutorials>.
- BARTSCH, H. J.: Matematické vzorce. Praha: SNTL, 1963. 575 s.
- HENDRY, D. F.: Econometrics – Alchemy or science? Oxford University Press, 2000. 542 s. ISBN 0-19-82954-2.
- HUŠEK, R., MAŇAS, M.: Matematické modely v ekonomii. Praha: SNTL, 1989. 402 s. ISBN 80-03-00098-X.
- HUŠEK, R.: Ekonometrická analýza. Praha: Ekopress, 1999. 304 s. ISBN 80-86119-19-X.
- KUENNE, R. E.: Readings in applied microeconomic theory. Oxford: Blackwell Publishers, 2000. 445 s. ISBN 0-631-22070-4.
- LANCASTER, K.: Modern consumer theory. Vermont: Edward Elgar, 1991. 242 s. ISBN 1-85278-384-2.
- LIMPOUCH, J.: Hledání extrémů funkcí. [on-line]. [cit. 12.12.2003]. Dostupné na <http://www-troja.fjfi.cvut.cz/limpouch/numet/extrem/ext1.html>.
- OBÍTKO, M.: Genetic algorithms. [on-line]. [cit. 12.12.2003]. Dostupné na <http://cs.felk.cvut.cz/xobitko/ga/main.html>.
- PESARAN, M. H., SCHMIDT, P.: Handbook of applied econometrics, volume II: Microeconomics. Oxford: Blackwell Publishers, 2000. 453 s. ISBN 0-631-21633-2.
- REKTORYS, K. a kol.: Přehled užití matematiky. Praha: SNTL, 1963. 1136 s.
- ROZSYPAL, S. a kol.: Přehled biologie. Praha: SPN, 1987. 686 s.
- SAS Institute, Inc.: The RELIABILITY procedure - Parameter estimation. [on-line]. [cit. 12.12.2003]. Dostupné na <http://certik.ruk.cuni.cz/sas/SASHTML/qc/chap30.sect34.htm>.
- SCHULTS, A. C.: GA archives. [on-line]. [cit. 12.12.2003]. Dostupné na <http://www.aic.nrl.navy.mil/galist/src>.
- SEAN, L.: Evolutionary computation systems. [on-line]. [cit. 12.12.2003]. Dostupné na <http://www.cs.emd.edu/projects/plus/ec/ecj/docs>.
- STUNDENMUND, A., H.: Using econometrics- a practical guide. London: The Addison Wesley, 2000. 637 s. ISBN 0-321-06481-X
- TIERNEY, L.: Maximalization and maximum likelihood estimation. [on-line]. [cit. 12.12.2003]. Dostupné na <http://www.stat.uiowa.edu/luke/xls/tutorial/techreport/nodes57.htm>.
- VERBEERK, M.: A guide to modern Econometric. Chichester: John Wiley&Sons, ltd., 2002. 386 s. ISBN 0-471-89982-8.

Výsledky uvedené v příspěvku jsou součástí řešení projektu č. 402/03/P130 „Vybrané aspekty postavení české ekonomiky v mezinárodním obchodu“ realizovaného za finanční podpory ze státních prostředků prostřednictvím Grantové agentury České republiky.

Adresa

Ing. Marcel Ševela, Ph.D., Ústav ekonomie, Mendelova zemědělská a lesnická univerzita v Brně, Zemědělská 1, 613 00 Brno, Česká republika