8. Prednáška

Evolučné stratégie

Historický úvod

1. Evolučná stratégia patrí historicky medzi prvé úspešné stochastické algoritmy. Bola navrhnutá už počiatkom 60-tych rokov Rechenbergom a Schwefelom

2. Evolučná stratégia vychádza zo všeobecných predstáv prirodzeného výberu, no o mnoho vágnejších ako napríklad u genetického algoritmu. Evolučná stratégia nie je založená na binárnej reprezentácii premenných, manipuluje priamo s "reálnou" reprezentáciou premenných.

$$\boldsymbol{x}_{opt} = \arg \min_{x \in [a,b]^n} f(\boldsymbol{x})$$

"**Personal story**"

Ingo Rechenberg, Hans-Paul Schwefel (teraz už obidva profesori) a Bienert sa začali v r. 1963 ako študenti na Technickej univerzite v Berlíne zaoberať experimentmi vo aerodynamický tuneli, kde optimalizovali tvar telies, aby tieto kládli čo najmenší odpor v prúdení. Pretože ani intuitívny prístup, ani jednoduché gradientové metódy neboli veľmi úspešné, začali sa zaoberať myšlienkou náhodných zmien parametrov definujúcich tvar, podobne ako je tomu u mutácií.

Evolučné stratégie sa od tej doby rozvinuli, no stále ostávajú spojené predovšetkým s technickými aplikáciami v inžinierskych oblastiach, a používajú sa najviac v Nemecku.

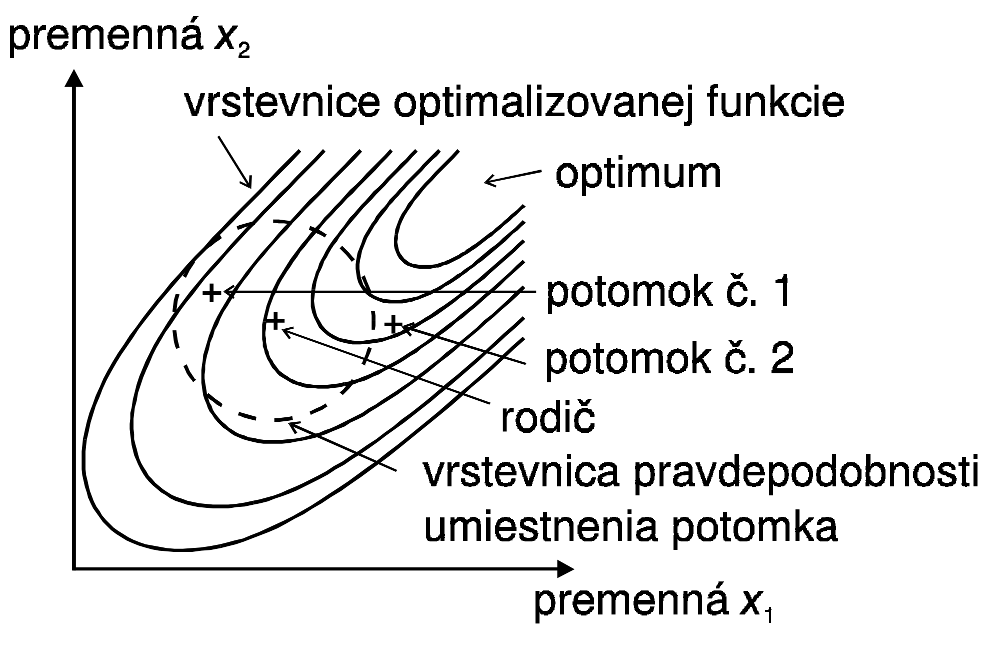
Základné princípy evolučnej stratégie

1. Základom evolučnej stratégie je nasledujúci predpis, ktorý "mutuje" aktuálne riešenie $\boldsymbol{x}$ na nové riešenie $\boldsymbol{x}'$

$$\boldsymbol{x}' = \boldsymbol{x} + \mathbf{N}(0,\sigma)$$

   kde $\mathbf{N}(0,\sigma)$ je vektor nezávislých náhodných čísel s normálnou distribúciou, nulovou strednou hodnotou a štandardnou odchýlkou σ.

2. Problém akceptácie nového riešenia $\boldsymbol{x}'$ je striktne deterministický, riešenie $\boldsymbol{x}'$ je akceptované (úspešné), keď $f(\boldsymbol{x}') < f(\boldsymbol{x})$.



Stratégia (1+1)
jeden rodič a jeden potomok

1. Základná heuristika je (1+1) stratégia

jedného rodiča a jedného potomka

Rechenberg robil základné pokusy na dvoch základných funkciách

(1) Sférickej funkcii

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} x_i^2$$

(2) Koridorová funkcia

$$f(\boldsymbol{x}) = -\sum_{i=1}^{n} x_i$$

2. Zmena veľkosti “mutácie” - teda vzdialenosti potomka od rodiča, a na “zmenu tejto zmeny”. Štandardná odchýlka σ sa v priebehu evolučnej stratégie mení podľa **pravidla 1/5 úspešnosti**. Nech $\varphi(k)$ je koeficient úspešnosti definovaný ako pomer počtu úspešných mutácií v priebehu posledných k iterácií k počtu k iterácií, z ktorých bola úspešnosť meraná, potom

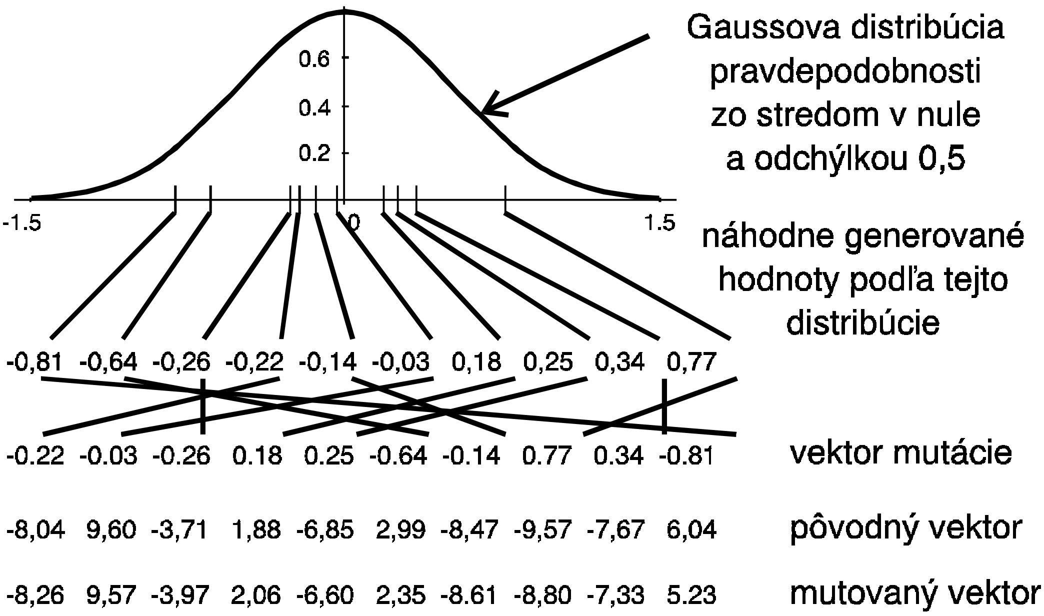
$$\sigma' = \begin{cases} c_d . \sigma \left(\varphi(k) < 1/5 \right) \\ c_i . \sigma \left(\varphi(k) > 1/5 \right) \\ \sigma \left(\varphi(k) = 1/5 \right) \end{cases}$$

kde $c_i>1$ a $c_d<1$ riadia zväčšovanie resp. zmenšovanie štandardnej odchýlky.

V literatúre sú tieto koeficienty špecifikované $c_d=0.82$ a $c_i=1/c_d=1.22$.

Intuitívne zdôvodnenie pravidla 1/5

Zvýšenie efektívnosti prehľadávania: pri úspešnosti by malo prehľadávanie pokračovať vo väčších krokoch, pri neúspešnosti by mali byť kroky kratšie. Toto pravidlo odvodené pre uvedené dve funkcie sa potom všeobecne používa aj pre ďalšie neznáme funkcie, pretože pre mnohé problémy jeho používanie pomáhalo udržať najrýchlejší postup k optimu.



Príklad mutácie vektora s reálnymi premennými za použitia Gaussovy distribúcie pravdepodobnosti pre určenie malej náhodnej.

Keď niektorá mutovaná hodnota x_i' nie je z intervalu povolených hodnôt, $x_i' \notin \langle a,b \rangle$, je potrebné použiť opravný proces, aby sa zaistilo, že opravená hodnota spadá do vymedzeného intervalu

$$x_i' := x_i + N(0,\sigma_i)$$

$$x_i' < a \Rightarrow x_i' := a$$
$$x_i' > b \Rightarrow x_i' := b$$

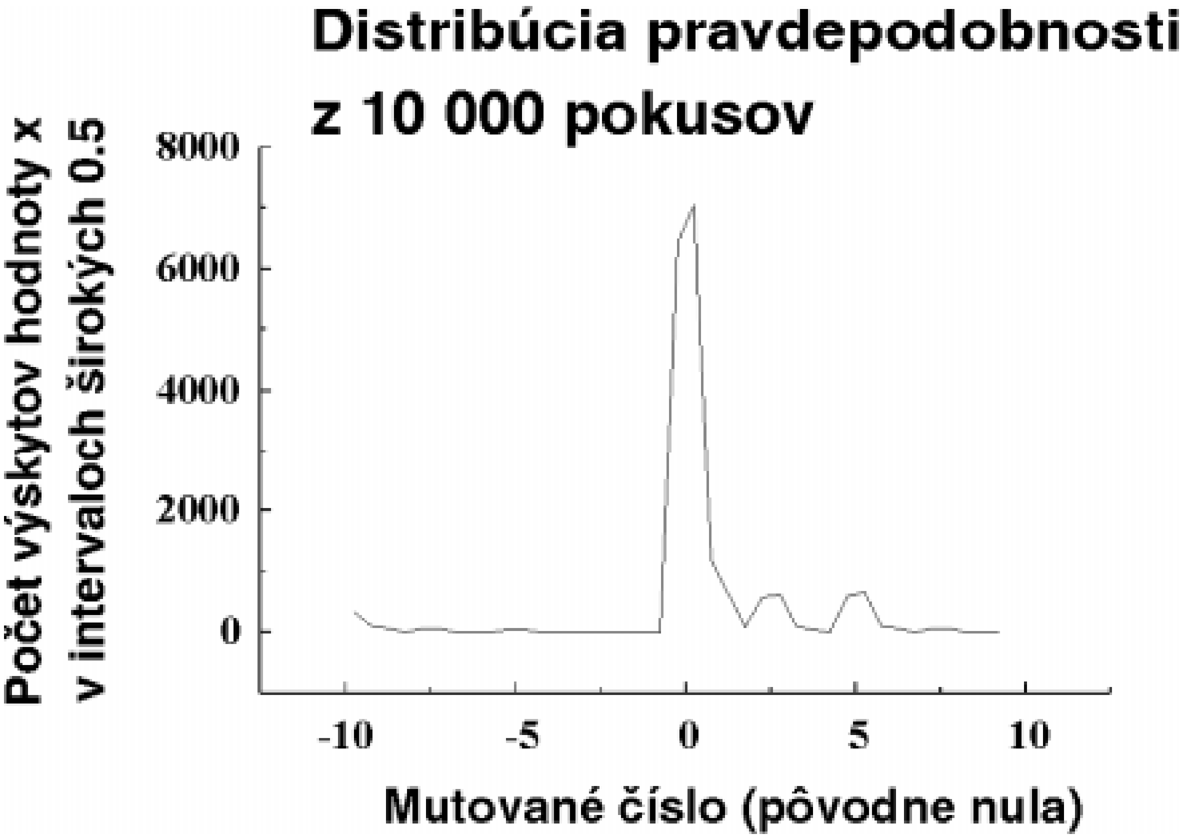
Pseudopascalovský kód
evolučná stratégia (1+1)

```
x:=náhodne generovaný vektor reálnych
   premenných;
t:=0; σ:=σini; x*:=x;
WHILE t<tmax DO
BEGIN i:=0; k:=0;
      WHILE i<imax DO
      BEGIN i:=i+1; x':=x+r(0,σ);
            IF f(x')<f(x) THEN
            BEGIN k:=k+1; x:=x';
                  IF f(x)<f(x*) THEN
                  x*:=x;
            END;
      END;
      IF k/imax<0.2 THEN σ:=cd*σ ELSE
      IF k/imax>0.2 THEN σ:=ci*σ;
END;
```

Generovanie náhodných čísel

Prečo je Gaussova distribúcia lepšia ako prehadzovania bitov v binárnom kóde?

000000000000 $\rightarrow$ -10.0
000000100000 $\rightarrow$ +0.5873
110000000000 $\rightarrow$ +5.00366
100000010000 $\rightarrow$ +0.805861
101000000000 $\rightarrow$ +2.50305
100000001000 $\rightarrow$ +0.41514
100100000000 $\rightarrow$ +1.25275
100000000100 $\rightarrow$ +0.21978
100010000000 $\rightarrow$ +0.27595
100000000010 $\rightarrow$ +0.1221
100001000000 $\rightarrow$ +0.15018
100000000001 $\rightarrow$ +0.0732601



Distribúcia pravdepodobnosti výsledku mutácie binárneho čísla 100000000000 je veľmi nerovnomerná, a dokonca nie je symetrická! Prečo? Dôvodom je tzv. **Hammingova bariéra**: binárne čísla 100000000000 a 011111111111 sú susedmi v reálnej reprezentácii, no bolo by potrebné zmeniť hodnoty všetkých 12 bitov, aby sme z prvého čísla dostali to druhé.

1. generátor náhodných čísel s normálnou distribúciou

```
help:=0;

FUNCTION Gauss1;
BEGIN  IF help=0 THEN
       BEGIN REPEAT v1:=2.0*random-1.0;
                    v2:=2.0*random-1.0;
                    r:= v1^2 + v2^2;
             UNTIL (r<1.0);
             fac:=sqrt(-2.0*ln(r)/r)
             savegaussdev:=v1*fac;
             Gauss1:=v2*fac;
             help:=1;
       END ELSE
       BEGIN help:=0;
             Gauss1:= savegaussdev;
       END
END
```

Generátor náhodnej premennej s normálnym (Gaussovským) rozdelením s nulovou strednou hodnotou a jednotkovou smerodajnou odchýlkou. Verzia podľa Numerical Recipes [12].

2. generátor náhodných čísel s normálnou distribúciou

```
help:=0;

FUNCTION Gauss2;
BEGIN  IF help=0 THEN
       BEGIN a:=√(-2.0*ln(random))
             b:=2*π*random;
             savegaussdev:=a*cos(b);
             Gauss2:=a*sin(b);
             help:=1;
       END ELSE
       BEGIN help:=0;
             Gauss2:= savegaussdev;
       END
END
```

Generátor náhodnej premennej s normálnym (Gaussovským) rozdelením s nulovou strednou hodnotou a jednotkovou smerodajnou odchýlkou. Jednoduchšia verzia (no vzhľadom k použitiu trigonometrických funkcií výpočtovo o niečo náročnejšia) používaná Schwefelom [4].

Generátor náhodných čísel s logistickou distribúciou

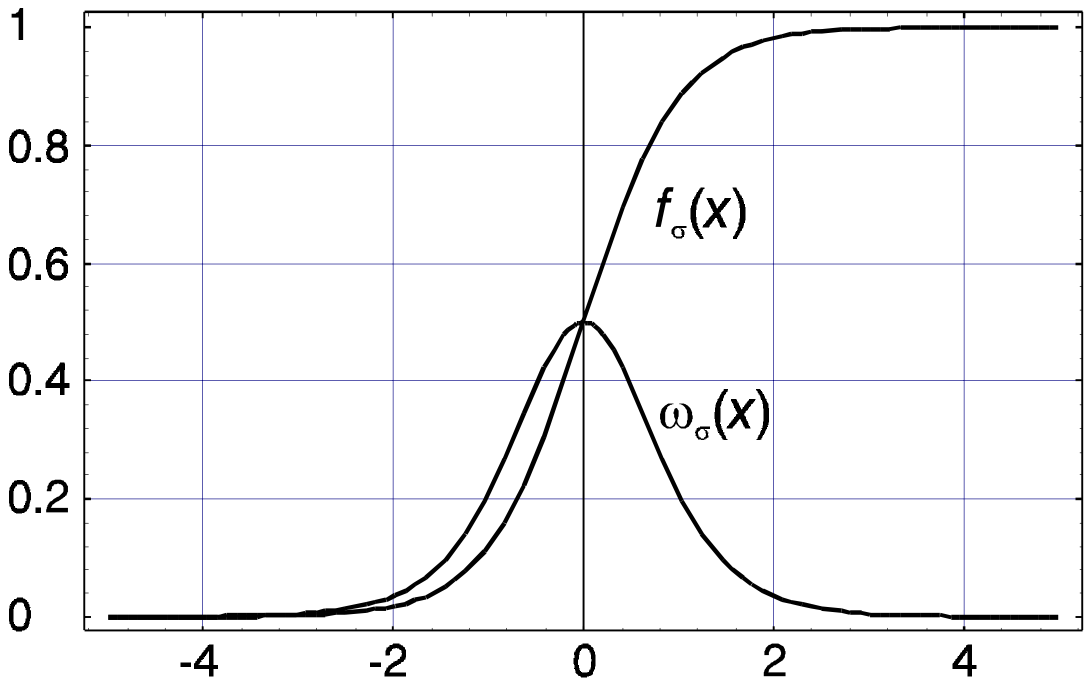
Pretože ale generátor náhodných čísel s normálnou distribúciou je používaný často a podstatne ovplyvňuje časovú náročnosť výpočtu, obrátime pozornosť na generátor náhodných čísel s **logistickou distribúciou**, ktorý je výpočtovo podstatne menej náročnejší ako predošlé dva generátory

$$f_\sigma(x) = \frac{1}{1+e^{-x/\sigma}}$$

$$\omega_\sigma(x) = f'(x) = \frac{1}{\sigma}\frac{e^{-x/\sigma}}{\left(1+e^{-x/\sigma}\right)^2}$$

kde parameter σ s kladnou hodnotou (približne odpovedajúci štandardnej odchýlke u normálnej distribúcie) určuje "šírku" hustotnej funkcie $\omega_\sigma(x)$.

Distribúcia $f_\sigma(x)$ mapuje celú reálnu os na otvorený interval (0,1), f_σ: R $\rightarrow$ (0,1).



Zobrazenie logistickej distribučnej funkcie $f_\sigma(x)$ a jej hustotnej funkcie $\omega_\sigma(x)$ pre $\sigma = 0$.

Logistická distribúcia bude použitá pre generovanie náhodných čísel

Keď vyriešime rovnicu

$$r = \frac{1}{1+e^{-x/\sigma}}$$

pre $r \in (0,1)$ dostávame

$$\boxed{x = \sigma \, ln \frac{r}{1-r}} \qquad (*)$$

Ak r je náhodné číslo s uniformnou distribúciou, potom x určené pomocou (*) je náhodné číslo s nulovou strednou hodnotou, splňujúce logistickú distribúciu.

Všeobecná teória konvergencie ku globálnemu minimu?

- Podobne, ako pre simulované žíhanie, bolo dokázané aj pre evolučnú stratégiu, že potenciálne poskytuje globálny extrém optimalizovanej funkcie $f(\mathbf{x})$.

- Za predpokladu regulárnosti optimalizovaného problému je možné dokázať konvergenčný teorém.

- Tento konvergenčný teorém tvrdí, že globálne optimum bude dosiahnuté s jednotkovou pravdepodobnosťou; nehovorí ale, za ak dlho generácií.

Viacčlenné stratégie

Schwefelom so spolupracovníkmi boli navrhnuté ďalšie sofistikovanejšie verzie evolučnej stratégie, takže v súčasnosti je možné už hovoriť o celej triede evolučných stratégií.

Schwefel novšie zaviedol viacčlenné evolučné stratégie, označované (m+l) alebo (m,l), ktoré imitujú ďalšie základné princípy biologickej evolúcie. Pracuje sa tu potom s celým súborom rodičovských vektorov $\boldsymbol{x}$, ktorých počet je m.

PLUS stratégia (m+l) vygenerujeme z rodičov l potomkov. Potom takto navrhnutých vektorov - potomkov a pôvodných “rodičovských” vektorov dovedna zoradíme podľa optimálnosti riešení, a prvých m najlepších jedincov - riešení zoberieme ako rodičov do ďalšej populácie.

“ČIARKA” stratégia (m,l) vyberáme budúcich m rodičov iba z momentálne vygenerovaných l potomkov, a rodičia týchto potomkov vymierajú.

Život každého jedinca je obmedzený iba na jednu generáciu.

Okrem mutácie je u obidvoch stratégií používané kríženie, teda čiastočná výmena informácií medzi vektormi reálnych čísel.

Kríženie dvoch alebo viacerých rodičov vždy produkuje jedného potomka, na rozdiel od genetických algoritmov, kedy dvaja rodičia generujú obvykle dvoch potomkov. Rovnako ako u genetických algoritmov môže byť chromozóm vybraný ako rodič aj niekoľkokrát.

$$\boldsymbol{y}_1 = a\,\boldsymbol{x}_1 + (1-a)\,\boldsymbol{x}_2$$
$$\boldsymbol{y}_2 = (1-a)\,\boldsymbol{x}_1 + a\,\boldsymbol{x}_2$$