WAE

Jarosław Arabas

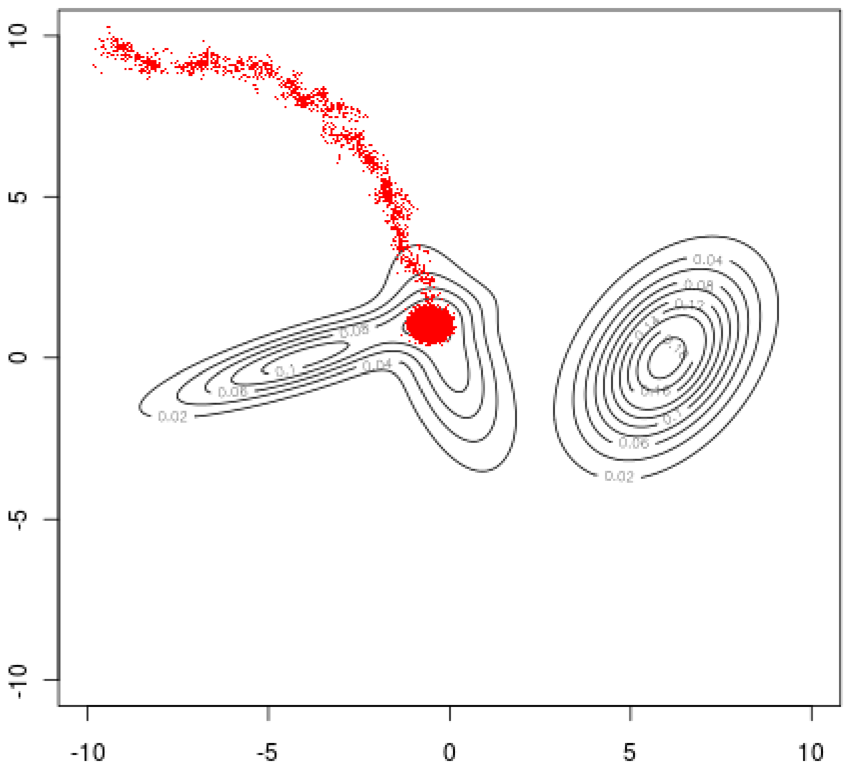
Adaptacja i samoczynna adaptacja parametrów AE

Algorytm CMA-ES

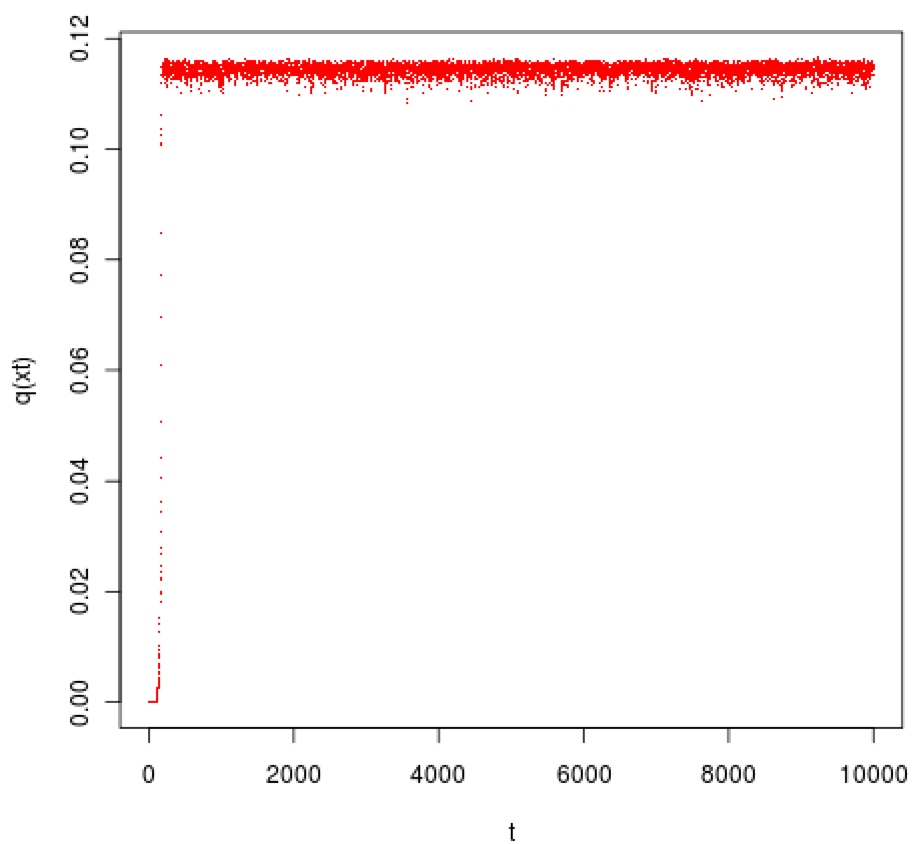
Dynamika mutacyjnego AE

- Mutacja gaussowska $\sigma = 0.1$

Wszystkie wygenerowane punkty



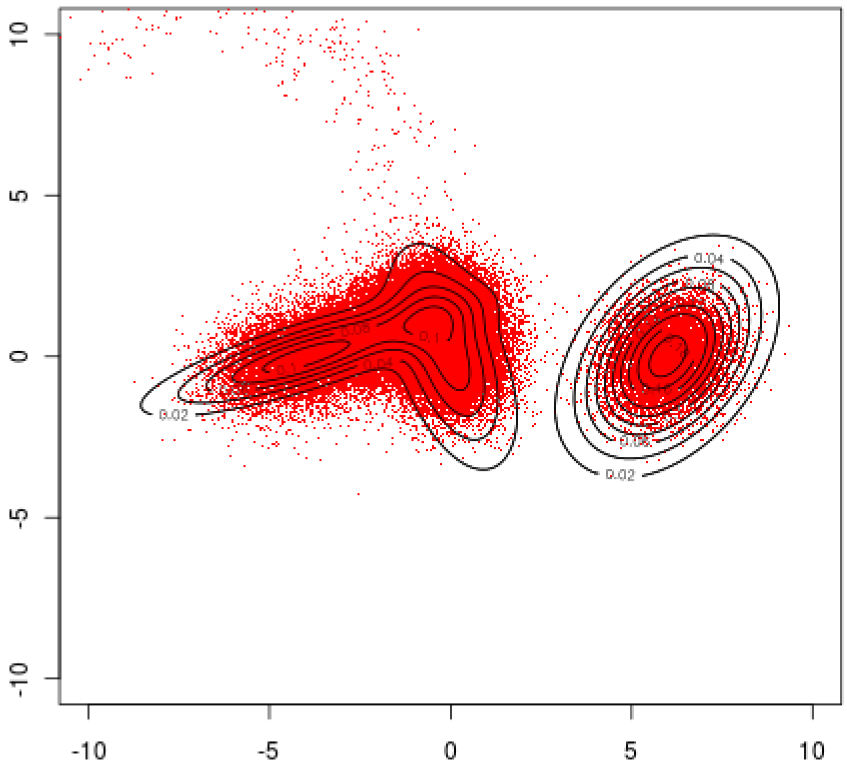
Wartość średnia jakości punktów populacji



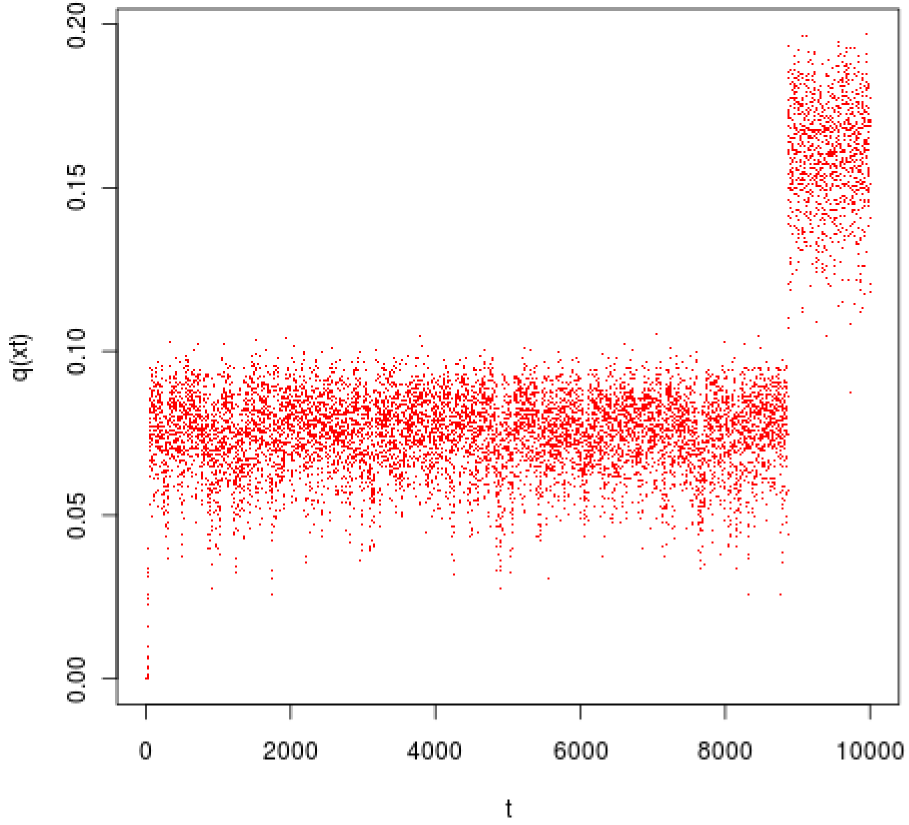
Dynamika mutacyjnego AE

- Mutacja gaussowska $\sigma = 0.55$

Wszystkie wygenerowane punkty



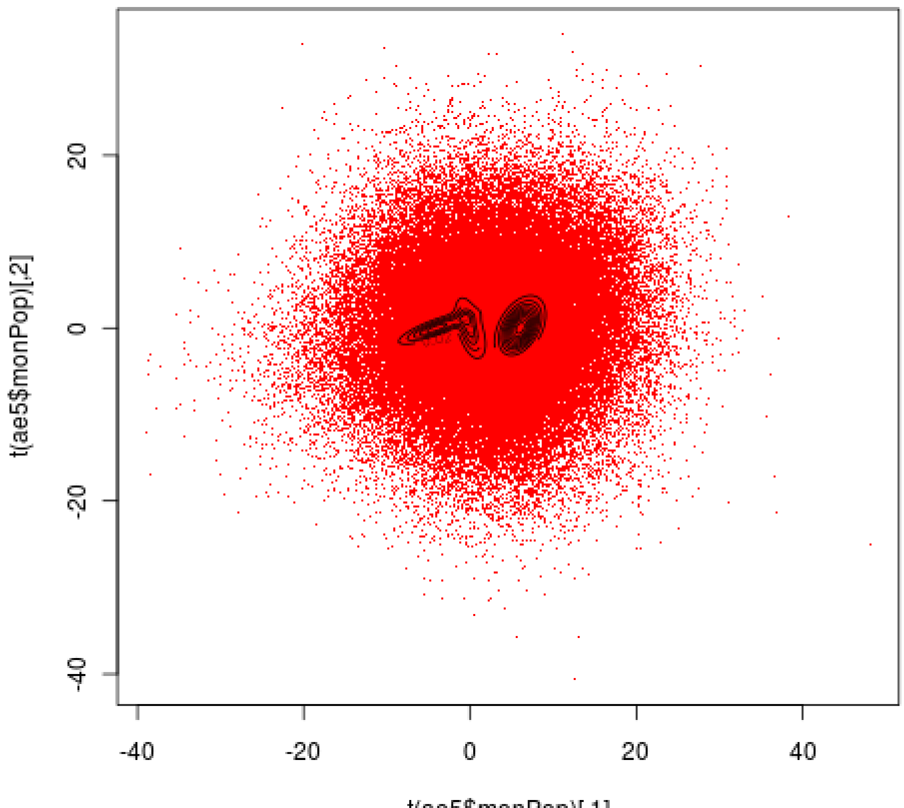
Wartość średnia jakości punktów populacji



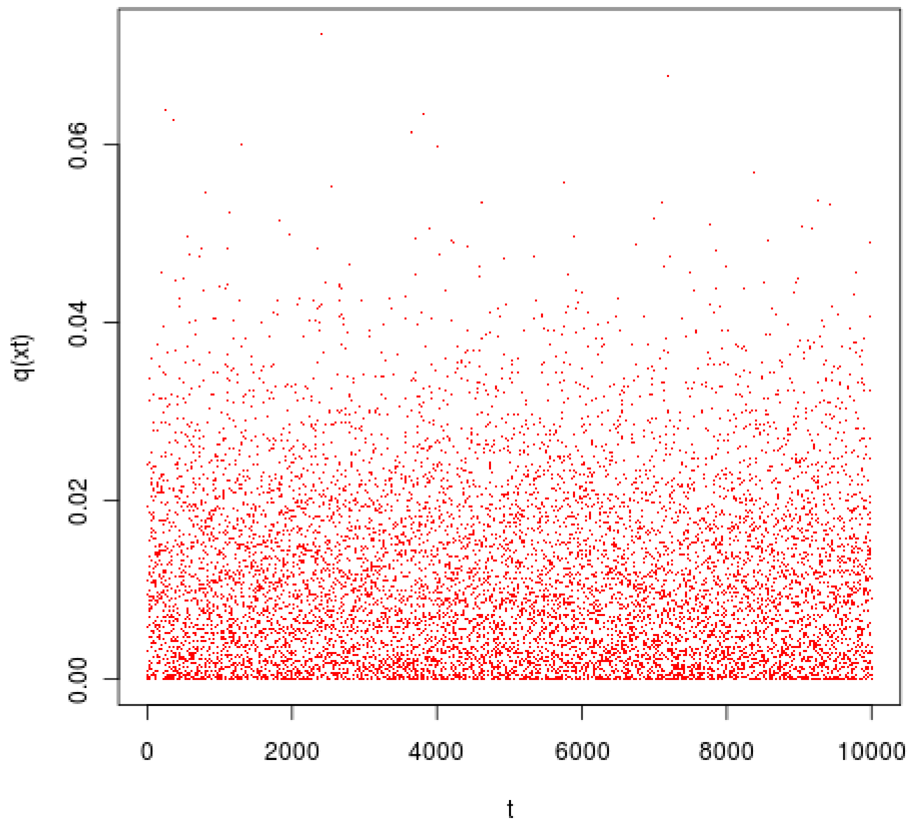
Dynamika mutacyjnego AE

- Mutacja gaussowska $\sigma=5$

Wszystkie wygenerowane punkty



Wartość średnia jakości punktów populacji



Eksploracja i eksploatacja

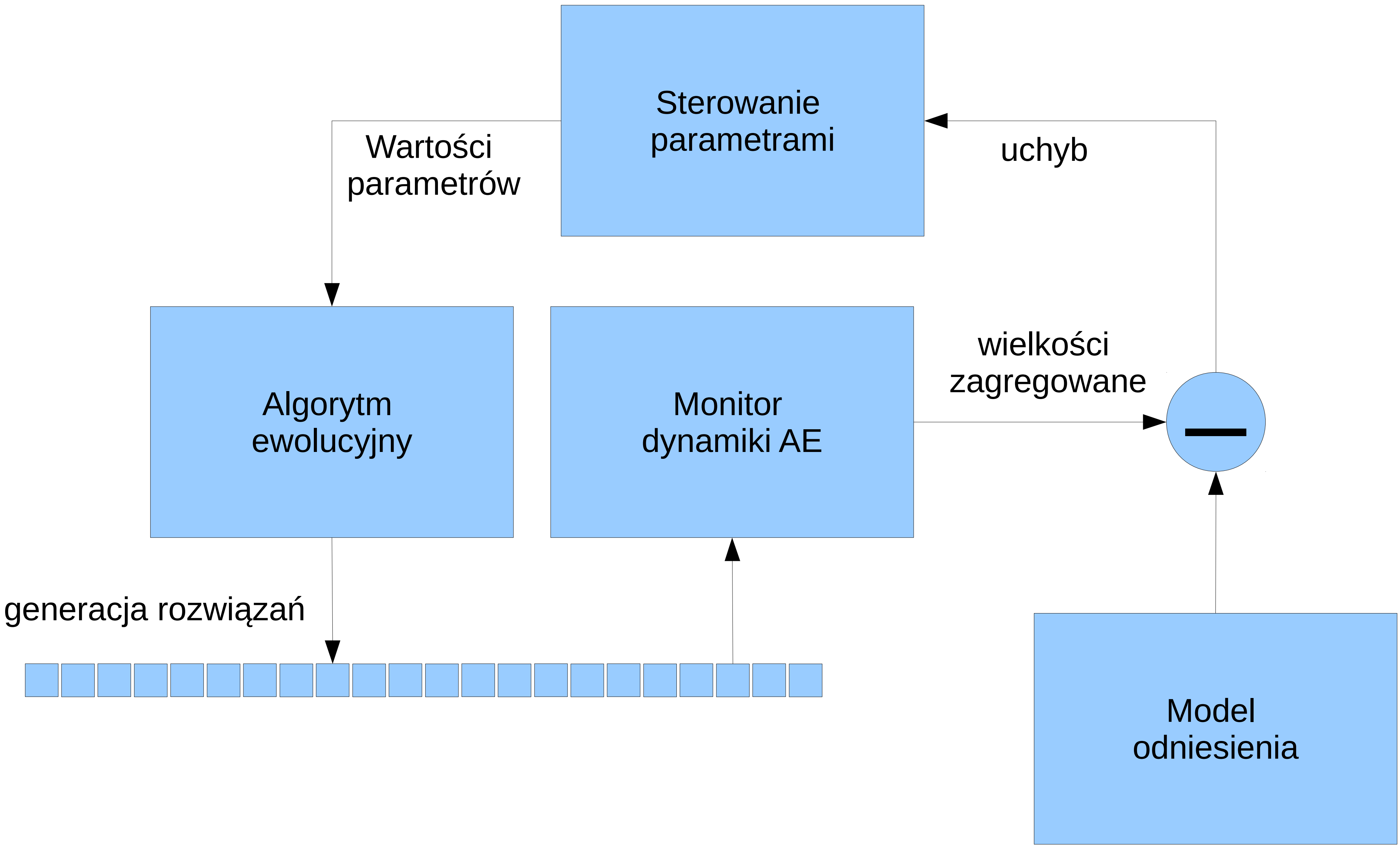
- Sterowanie presją selekcji
  - Reprodukcja
    - progowa – wartość θ
    - turniejowa – wielkość szranek s
    - proporcjonalna – modyfikacja wartości funkcji celu (*fitness scaling*)

      Im większe zróżnicowanie prawdopodobieństwa selekcji, tym większa presja selekcji
  - Sukcesja – jeśli jest elitarna, to zwiększa presję selekcji
- Rozłożenie populacji w przestrzeni zależne od funkcji celu

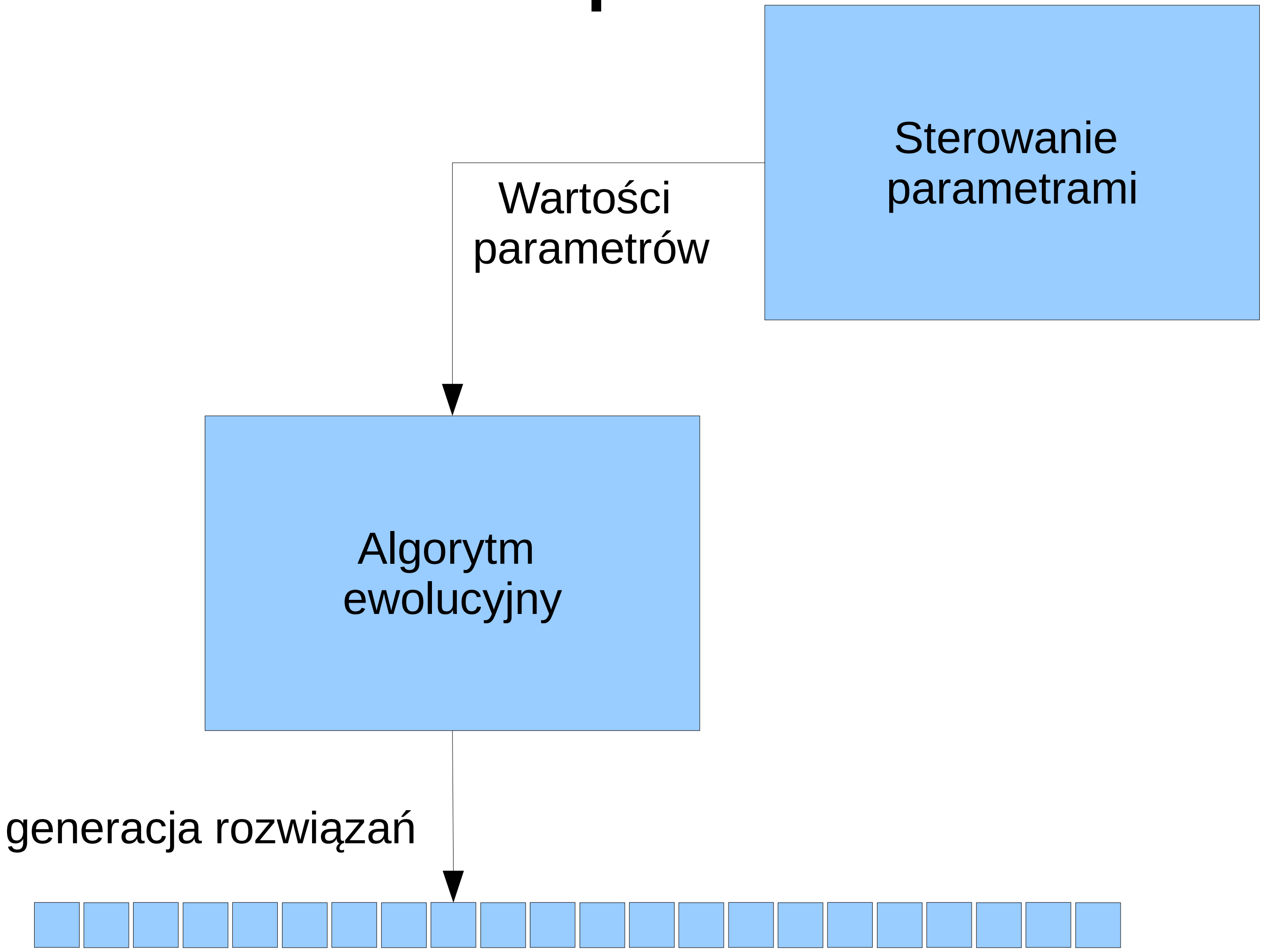
Eksploracja i eksploatacja

- Rozpraszanie populacji w sposób niezależny od wartości funkcji celu
  - Mutacja - zasięg mutacji (wariancja v_m) - im jest większa, tym większa różnorodność
  - Krzyżowanie uśredniające – im większe jego prawdopodobieństwo p_c , tym mniejsza różnorodność
  - Krzyżowanie wymieniające – rozprasza populację poprzez częściową dekorelację wymiarów

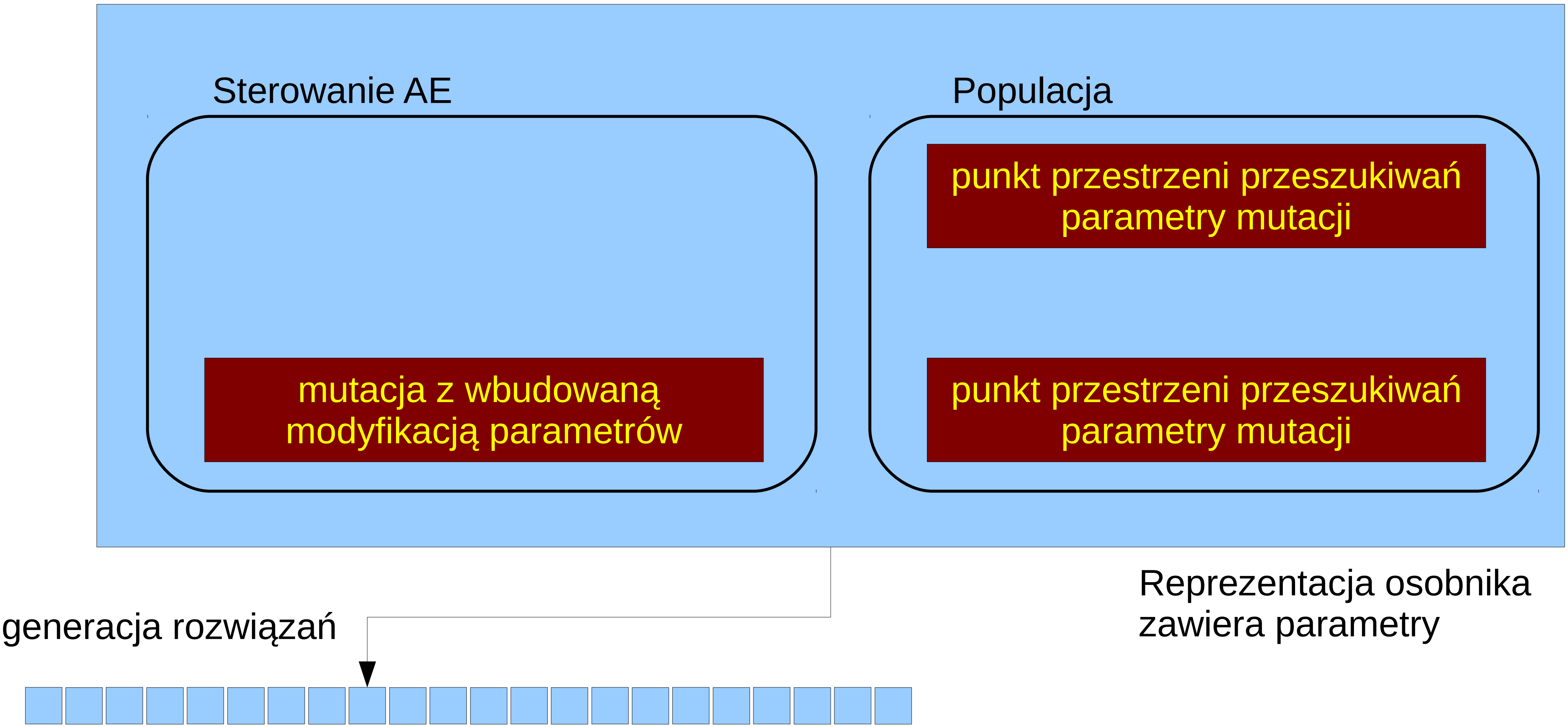
Adaptacja parametrów AE



Zaprogramowane sterowanie parametrami AE



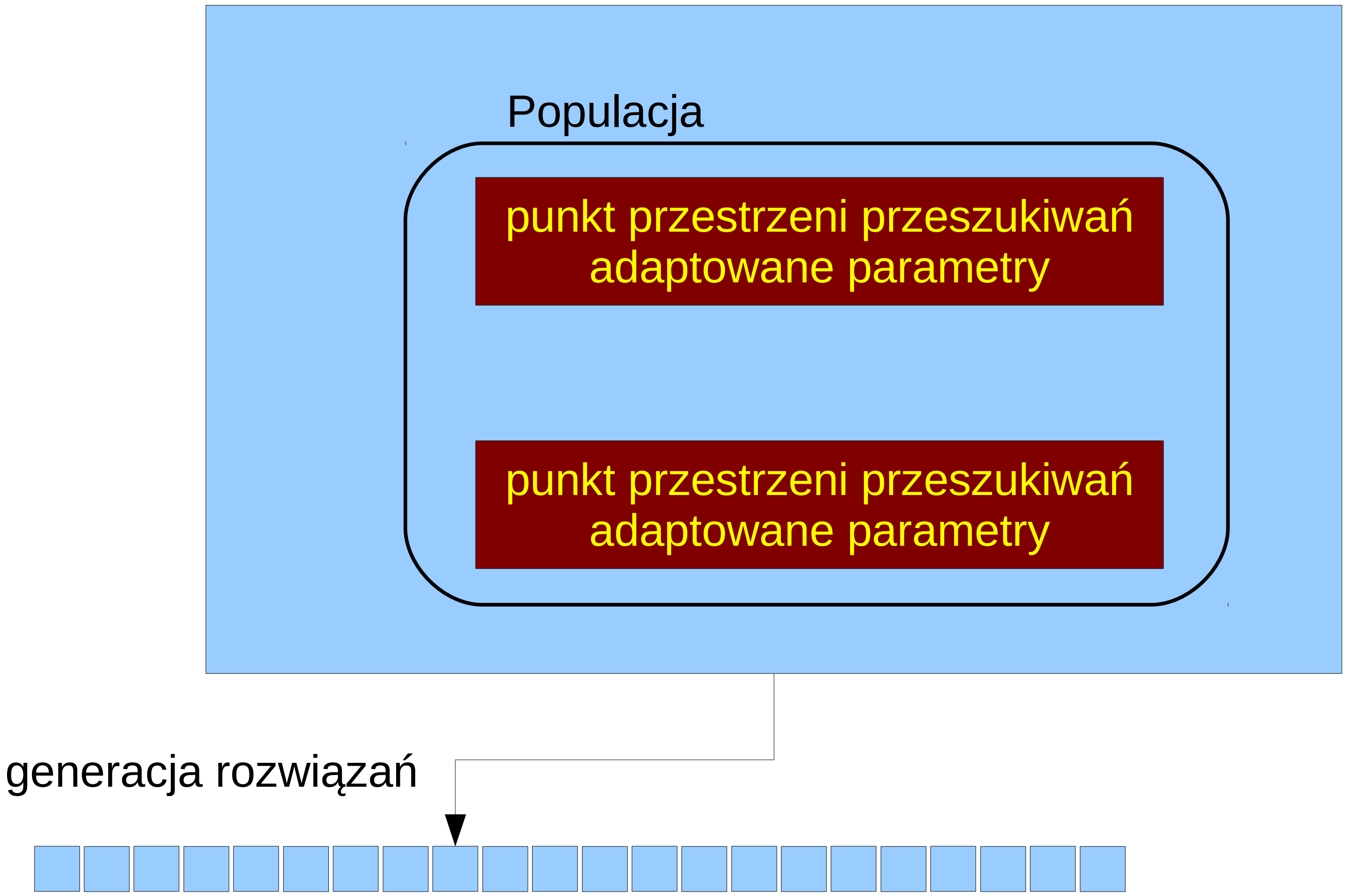
Samoczynna adaptacja parametrów AE



Reprezentacja osobnika zawiera parametry

Parametry określają sposób przekształcania punktu

Samoczynna adaptacja parametrów AE



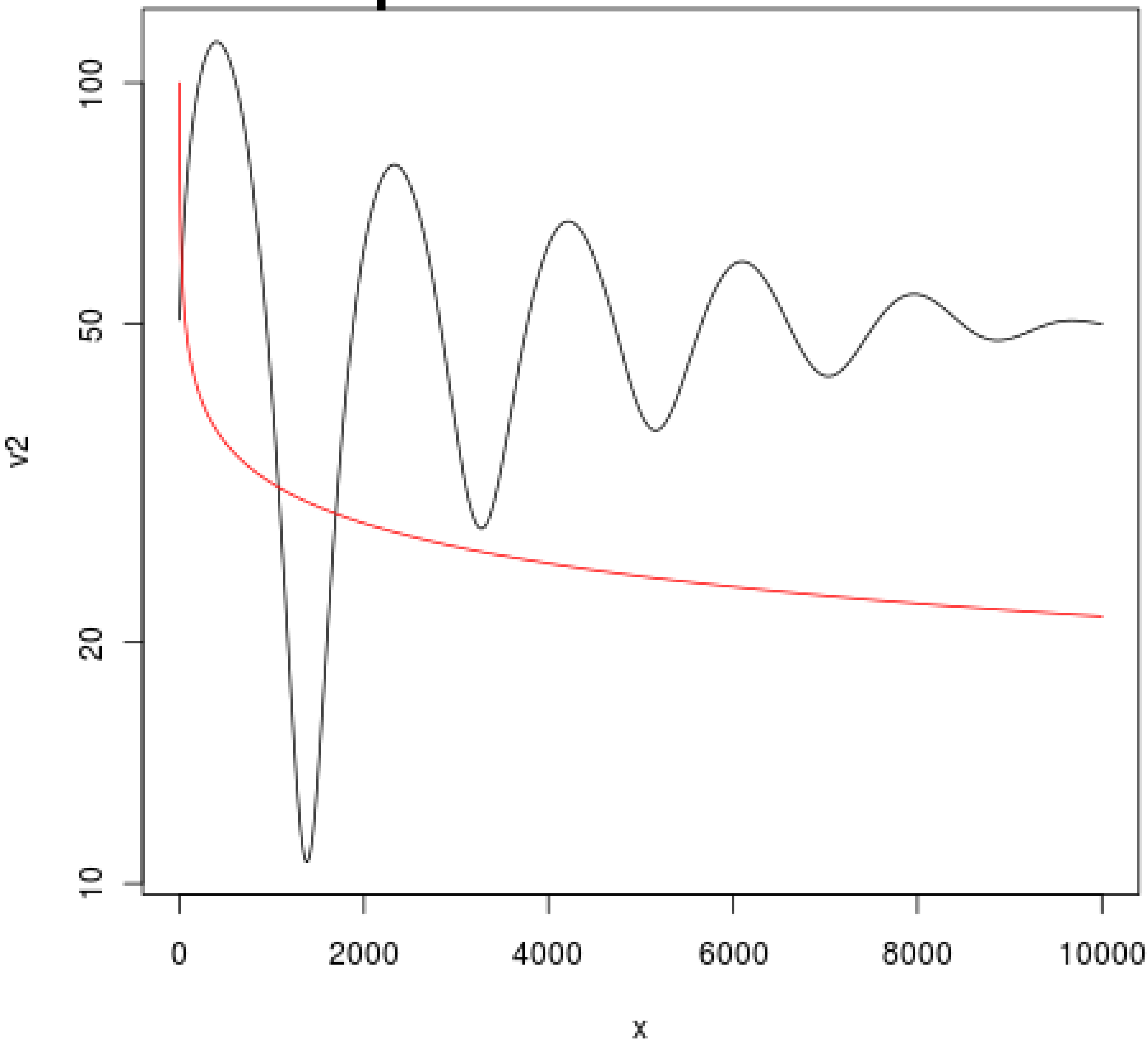
Reprezentacja osobnika zawiera parametry

Parametry określają sposób przekształcania punktu

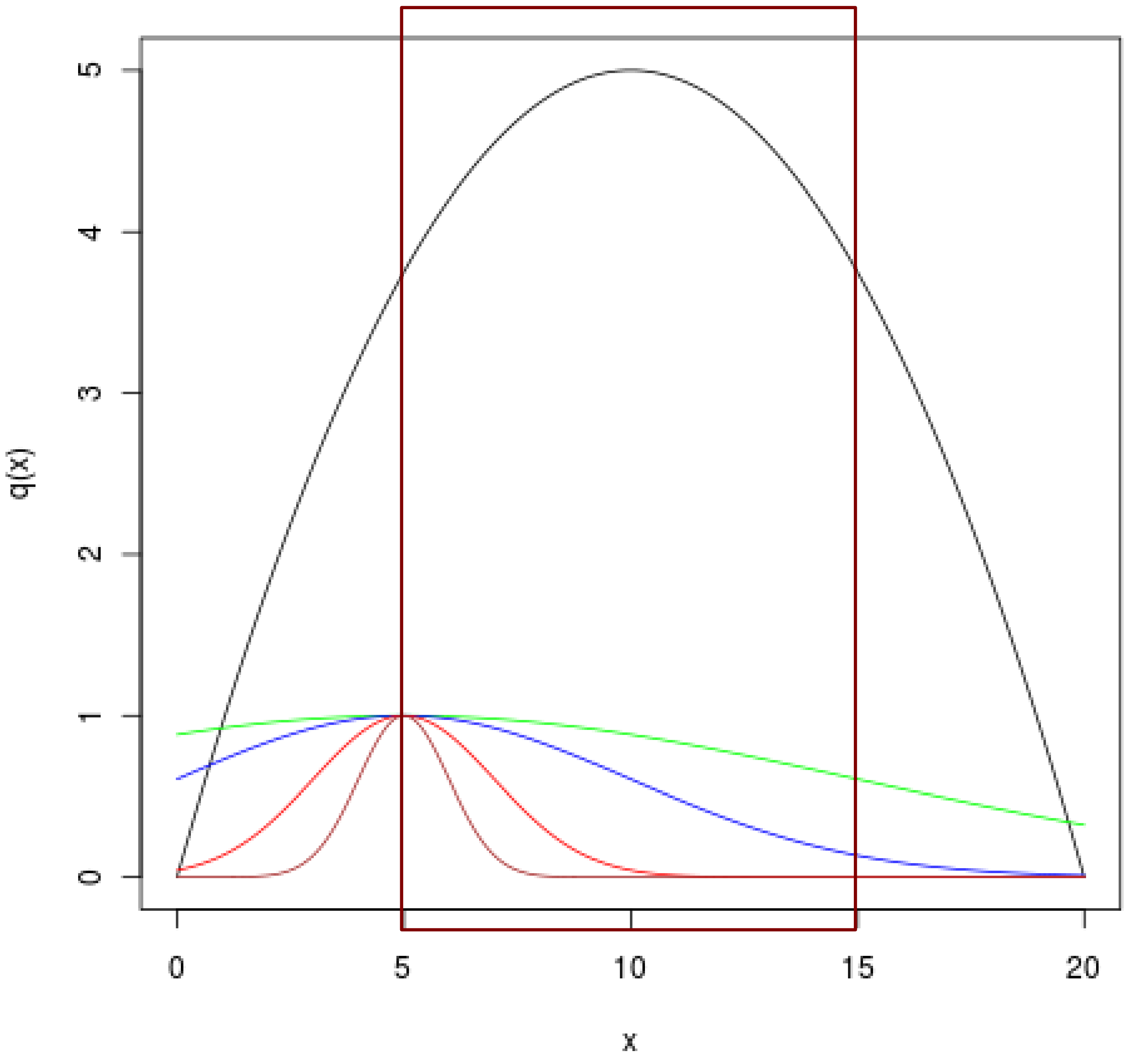
Przykład adaptacji mutacji

- Sterowanie zaprogramowane
- Reguła 1/5 liczby sukcesów
- Samoczynna adaptacja (Schwefel, Rechenberg)
- adaptacja parametrów

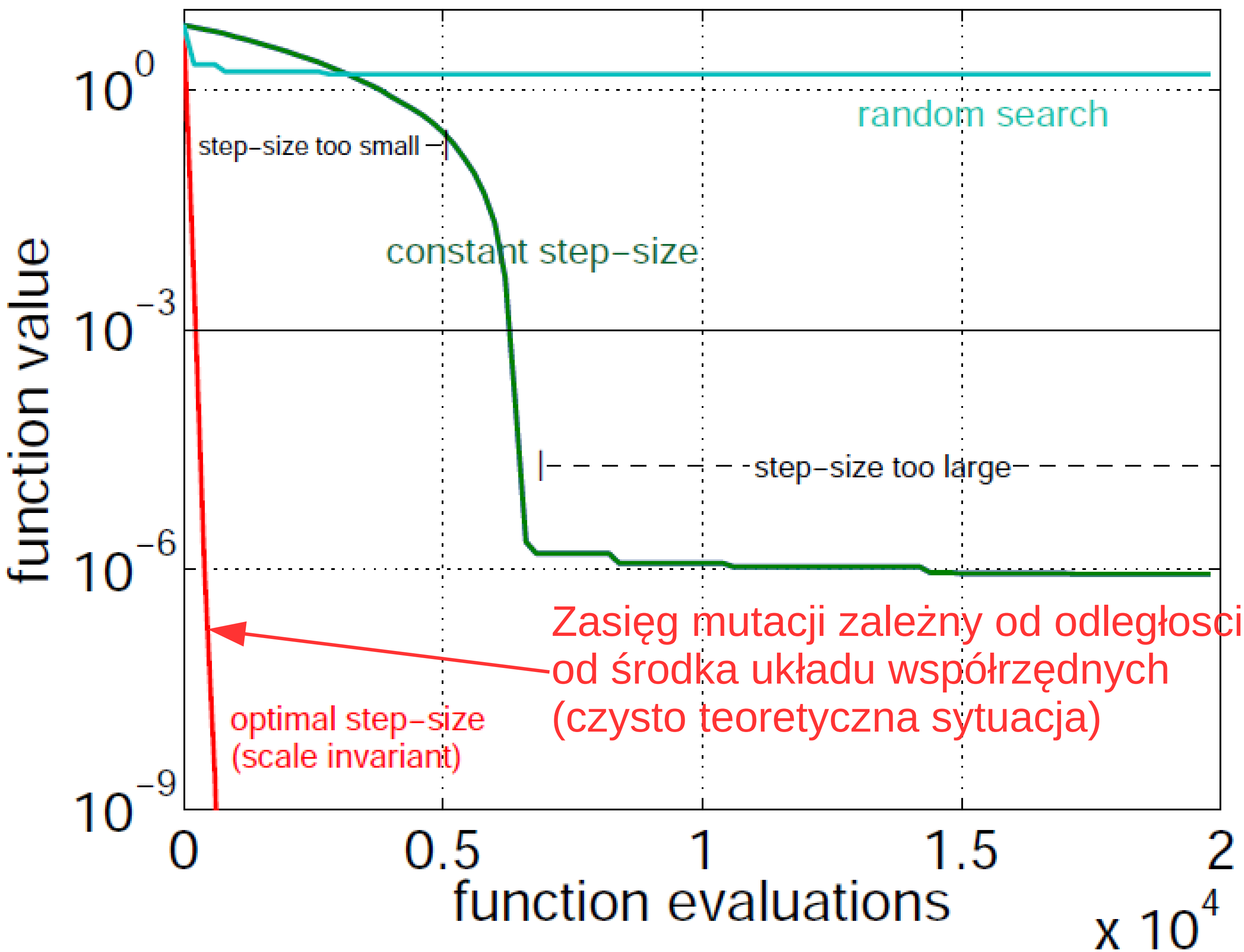
Zaprogramowane sterowanie parametrami



Liczba sukcesów w wyniku mutacji



Zależność tempa zbieżności od zasięgu mutacji (alg. wspinaczkowy)

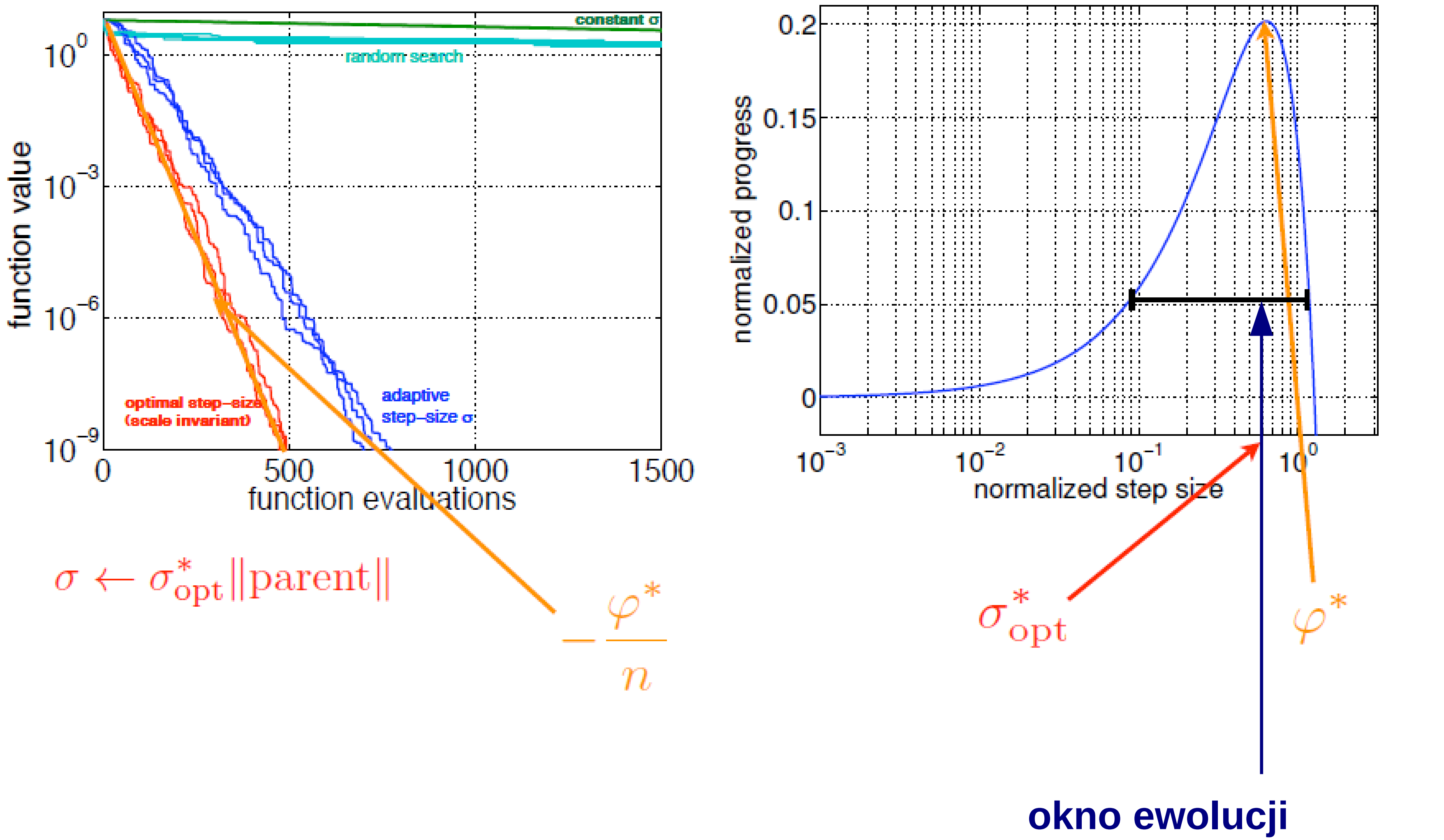


$$f(\boldsymbol{x}) = \sum_{i=1}^{n} x_i^2$$

in $[-0.2, 0.8]^n$
for $n = 10$

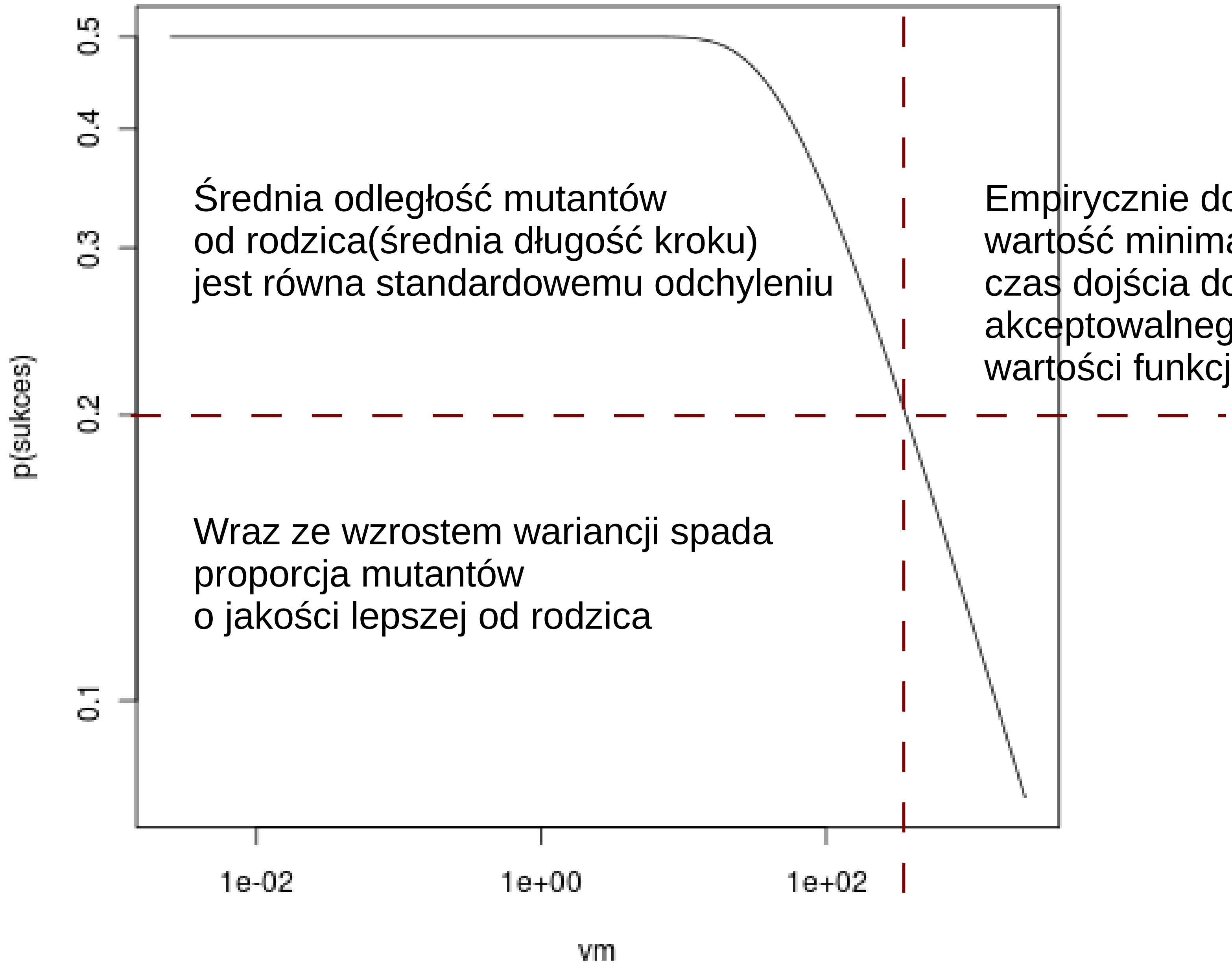
Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Zależność tempa zbieżności od zasięgu mutacji (alg. wspinaczkowy)



Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Reguła 1/5 liczby sukcesów



Średnia odległość mutantów
od rodzica(średnia długość kroku)
jest równa standardowemu odchyleniu

Empirycznie dobrana
wartość minimalizująca
czas dojścia do
akceptowalnego poziomu
wartości funkcji celu

Wraz ze wzrostem wariancji spada
proporcja mutantów
o jakości lepszej od rodzica

Reguła 1/5 liczby sukcesów

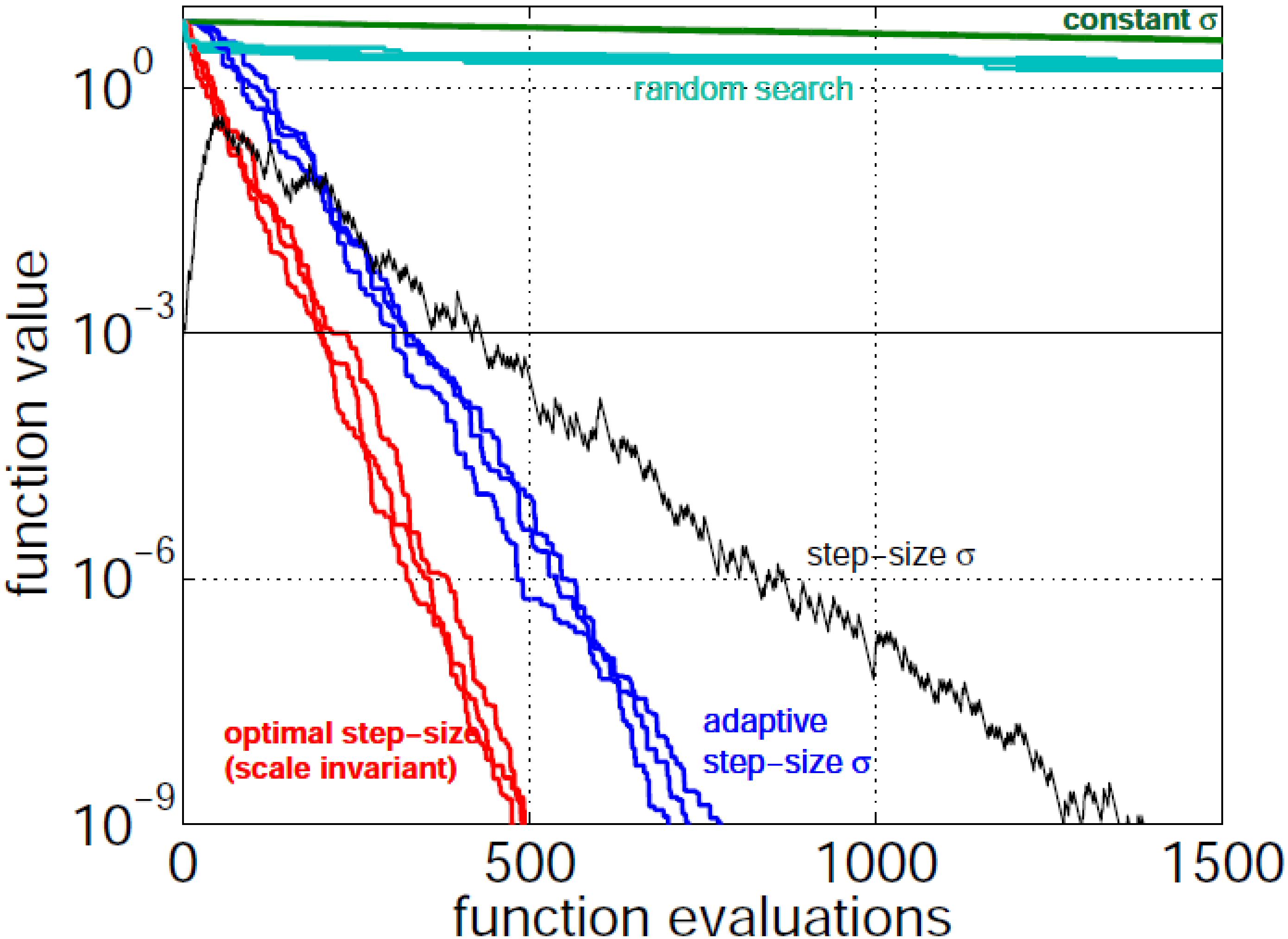
- Liczba mutacji w wyniku których mutant jest lepszy od rodzica powinna wynosić 1/5
- Sterowanie wariancją mutacji
- Pierwotnie stosowana do (1,lambda)-ES, czyli algorytmu wspinaczkowego

$$\sigma(t+k)=a\sigma(t) \quad gdy \quad p(sukces)<0.2$$

$$\sigma(t+k)=\frac{1}{a}\sigma(t) \quad gdy \quad p(sukces)>0.2$$

$$0.817\leq a\leq 1$$

Reguła 1/5 liczby sukcesów

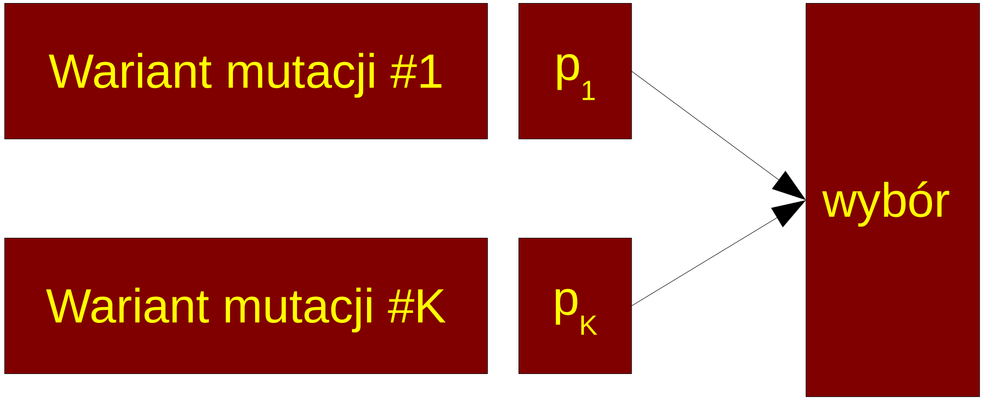


$$f(\boldsymbol{x}) = \sum_{i=1}^{n} x_i^2$$

in $[-0.2, 0.8]^n$
for $n = 10$

Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Mutacja wariantowa z wyborem zależnym od poprawy



Prawdopodobieństwo wyboru zależne od tego, ile mutantów było lepsze od ich rodziców

Zapominanie

Mutacja wariantowa z wyborem zależnym od poprawy

$$
\begin{aligned}
&\textbf{procedure}\ mutation(x)\\
&\quad j^* = select\ j\\
&\qquad\qquad where\ p_{sel}(j) = \frac{n_s(j)}{\sum_{k=1}^{K} n_s(k)}\\
&\quad y = mutation(x, j^*)\\
&\quad if\ (success(j^*))\\
&\qquad n_s(j^*) = n_s(j^*) + \alpha\\
&\quad n_s(j) = \beta \cdot n_s(j)\\
&\quad return(y)
\end{aligned}
$$

ns – skumulowana liczba sukcesów dla każdego wariantu mutacji

Samoczynna adaptacja zasięgu mutacji

$$\textbf{procedure}\ mutation(x)$$

$$\sigma_j = \sigma_j \exp(\tau a + \tau' a_j)$$

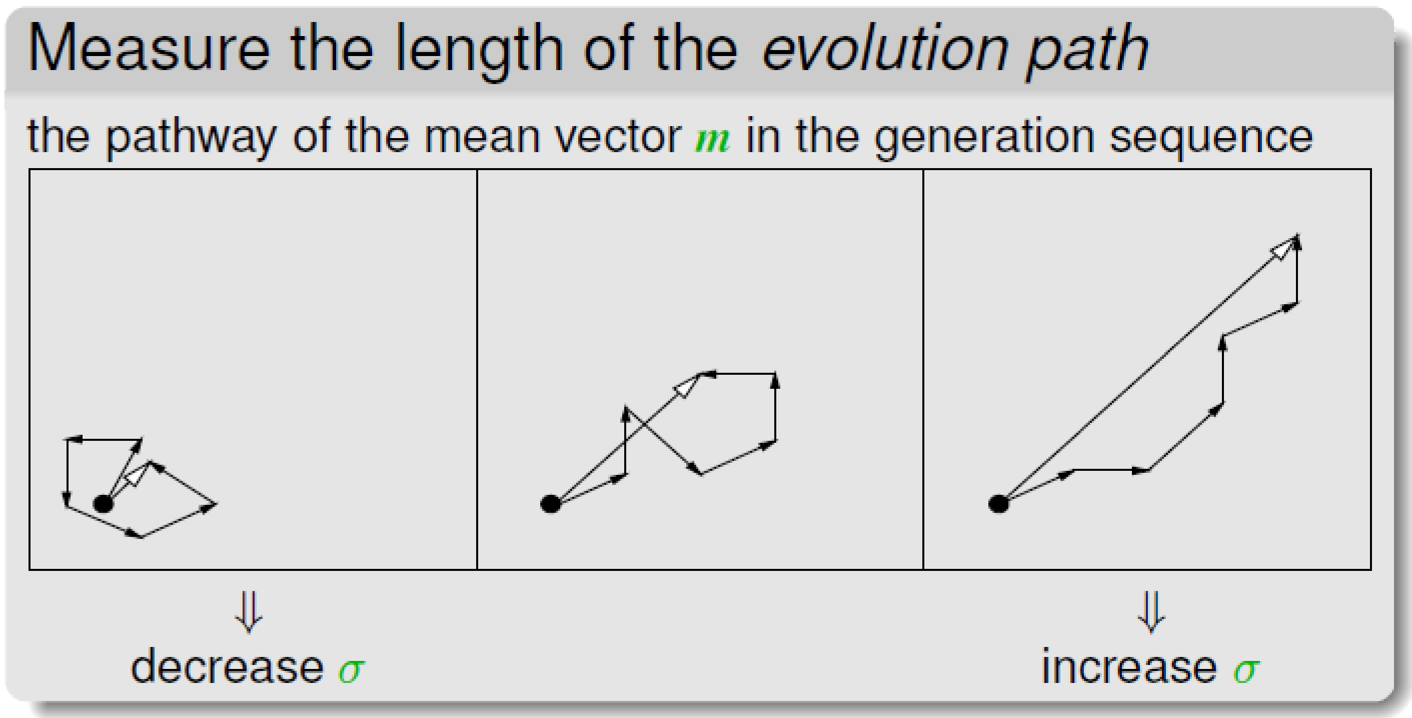
$$where\ \tau = \frac{1}{\sqrt{2\mathrm{n}}}, \tau' = \frac{1}{\sqrt{2\sqrt{n}}}, a \sim N(0,1), a_j \sim N(0,1)$$

$$y_j = x_j + \sigma_j d_j$$

$$where\ d_j \sim N(0,1)$$

$$return(y)$$

Adaptacja skumulowanego kroku algorytm CSA-ES



Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Adaptacja skumulowanego kroku algorytm CSA-ES

Przestrzeń x

Przestrzeń skojarzona

$$c_\sigma \approx 4/n, \quad d_\sigma \approx 1, \quad p_\sigma = 0$$

$$\textit{while}\,!\,\textit{stop}$$

$$\quad \textit{generuj}\ \boxed{d_i(t)} \sim N(0, I), i = 1 \ldots \lambda$$

$$\quad \textit{oblicz}\ \boxed{q_i(t)} = q\left(\boxed{m(t) + \sigma(t) \cdot d_i(t)}\right)$$

$$\quad \textit{sortuj według}\ q_i(t)$$

$$\quad \boxed{\Delta(t)} = \frac{1}{\mu} \sum_{i=1}^{\mu} d_i(t)$$

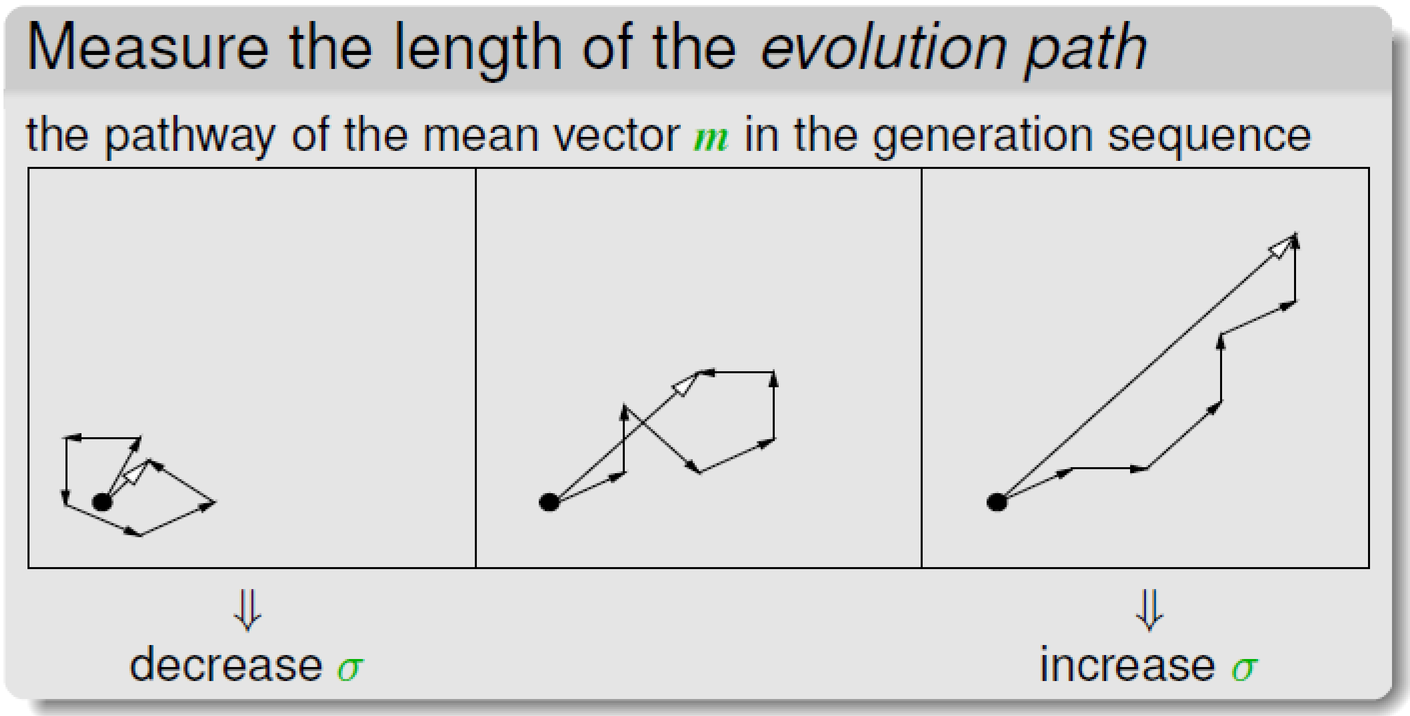
$$\quad m(t+1) = m(t) + \sigma(t) \cdot \Delta(t)$$

$$\quad \boxed{p_\sigma(t+1) = (1 - c_\sigma) p_\sigma(t) + \sqrt{1 - (1 - c_\sigma)^2} \sqrt{\mu}\, \Delta(t)}$$

$$\quad \boxed{\sigma(t+1)} = \sigma(t) \cdot \exp\left(\frac{c_\sigma}{d_\sigma}\left(\boxed{\frac{\|p_\sigma\|}{E\|N(0, I)\|}} - 1\right)\right)$$

$$\quad t \leftarrow t + 1$$

Adaptacja skumulowanego kroku algorytm CSA-ES



Trajektoria punktu środkowego
w przestrzeni skojarzonej

Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Modyfikacja macierzy kowariancji algorytm CMA-ES (wersja 0)

Przestrzeń x

$$C(1)=I$$

$$while\,!\,stop$$

$$\quad generuj\; d_i(t) \sim N(0, C(t)),\, i=1\ldots\lambda$$

$$\quad oblicz\; q_i(t) = q\big(m(t)+\sigma(t)\cdot d_i(t)\big)$$

$$\quad sortuj\; według\; q_i(t)$$

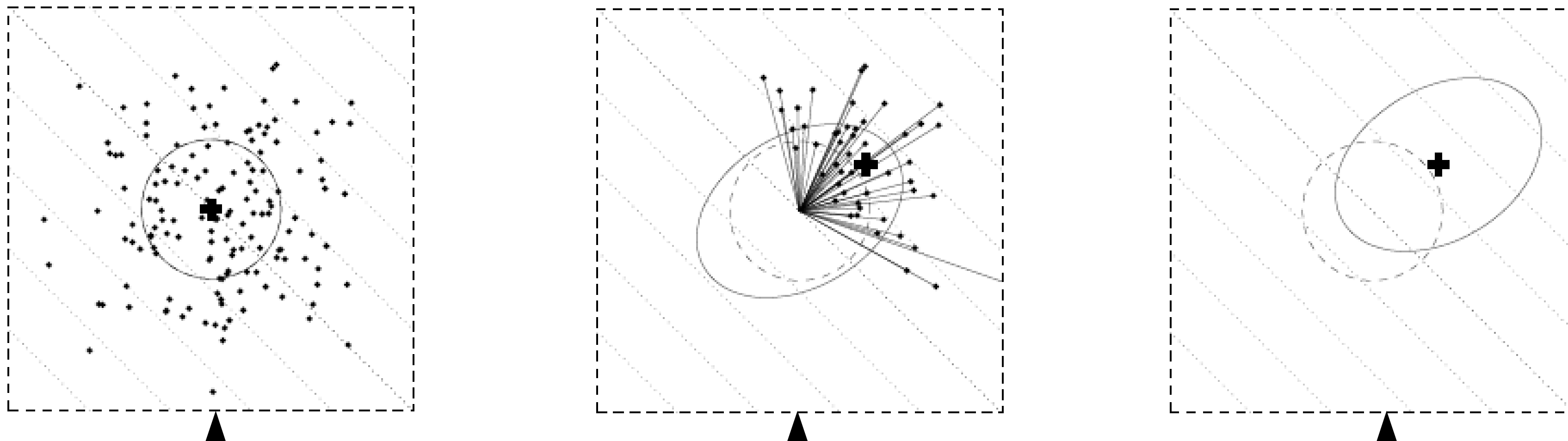
$$\quad \Delta(t) = \frac{1}{\mu}\sum_{i=1}^{\mu} d_i(t)$$

$$\quad \boxed{m}(t+1) = m(t)+\sigma(t)\cdot\Delta(t)$$

$$\quad \boxed{C}(t+1) = (1-c_{cov})C(t)+c_{cov}\,\mu\,\Delta(t)\Delta(t)^T$$

$$\quad t \leftarrow t+1$$

Adaptacja macierzy kowariancji algorytm CMA-ES



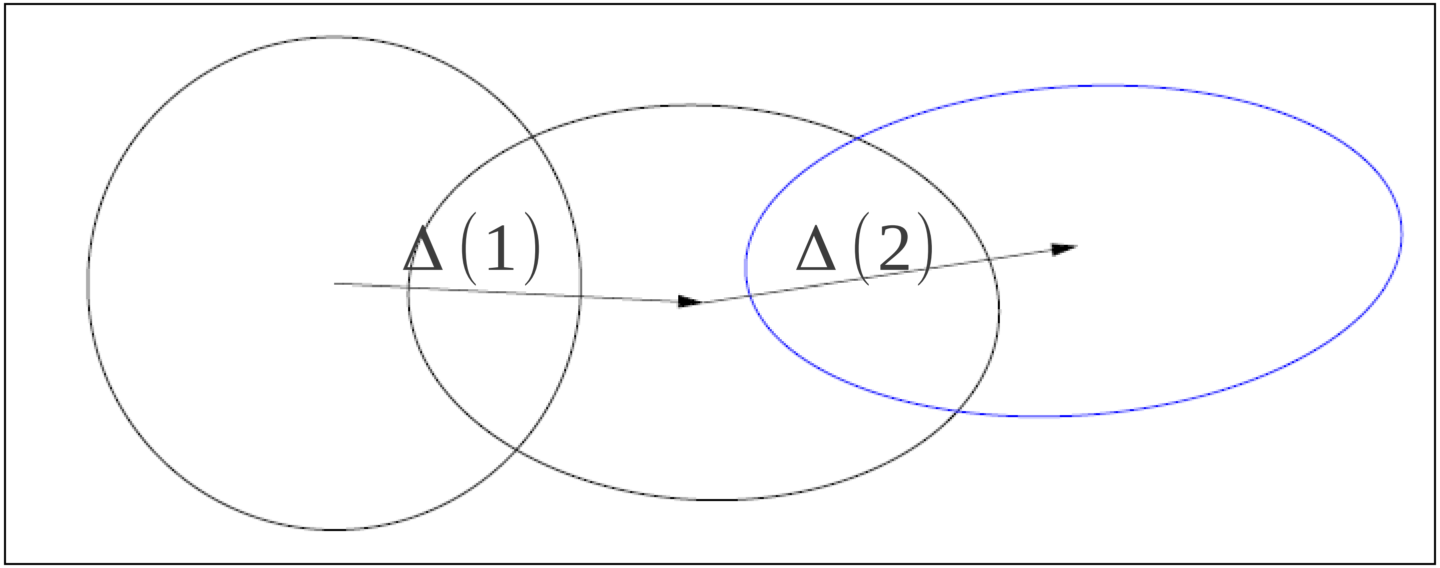
$$\Delta(t)=\frac{1}{\mu}\sum_{i=1}^{\mu} d_i(t)$$

$$d_i(t) \sim \sigma(t)\cdot N(0, C(t)), i=1\ldots\lambda$$

$$C(t+1)=(1-c_{cov})C(t)+c_{cov}\mu\,\Delta(t)\,\Delta(t)^T$$

Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Adaptacja macierzy kowariancji algorytm CMA-ES



Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Modyfikacja macierzy kowariancji algorytm CMA-ES (wersja 1)

Przestrzeń x

$$C(1)=I, \quad p_c(1)=0$$

$$\textit{while ! stop}$$

$$\quad \textit{generuj}\ d_i(t) \sim N(0, C(t)), i=1\ldots\lambda$$

$$\quad \textit{oblicz}\ q_i(t)=q\left(m(t)+\sigma(t)\cdot d_i(t)\right)$$

$$\quad \textit{sortuj według}\ q_i(t)$$

$$\quad \Delta(t)=\frac{1}{\mu}\sum_{i=1}^{\mu} d_i(t)$$

$$\quad \boxed{m}(t+1)=m(t)+\sigma(t)\cdot\Delta(t)$$

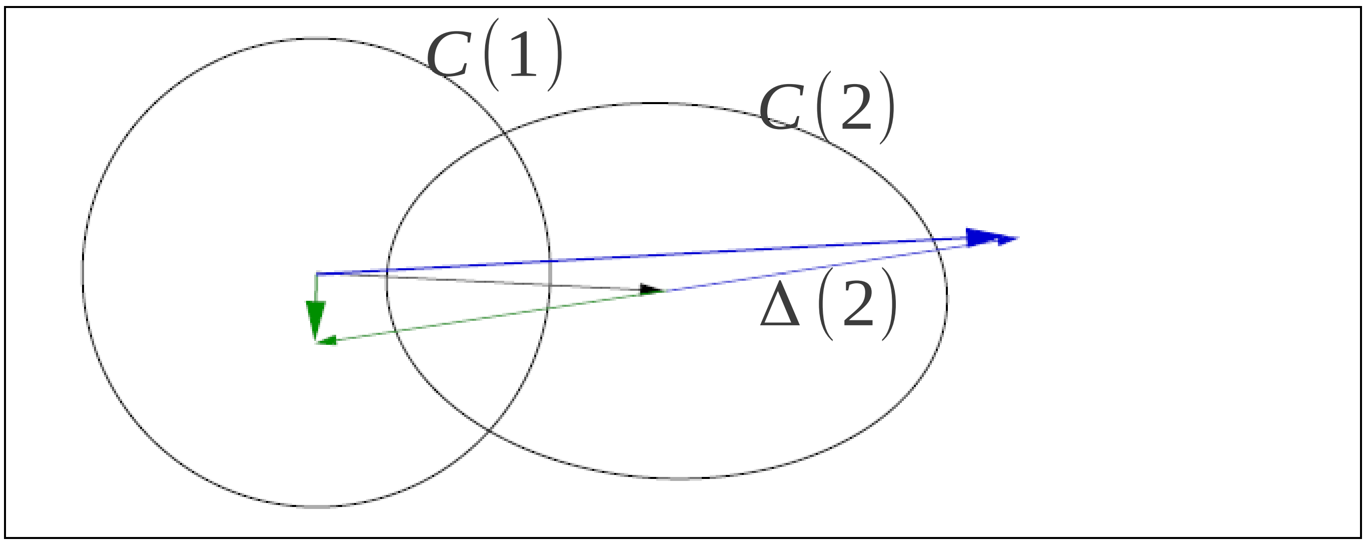
bezwładność

$$\quad \boxed{p_c}(t+1)=(1-c_c)\,p_c(t)+\sqrt{1-(1-c_c)^2}\,\sqrt{\mu}\,\Delta(t)$$

$$\quad \boxed{C}(t+1)=(1-c_{cov})\,C(t)+c_{cov}\,p_c(t+1)\,p_c(t+1)^T$$

$$\quad t \leftarrow t+1$$

Adaptacja skumulowanego kroku algorytm CMA-ES



$$p_c(t+1) = (1-c_c)\,p_c(t) + \sqrt{1-(1-c_c)^2}\,\sqrt{\mu}\,\Delta(t)$$
$$C(t+1) = (1-c_{cov})\,C(t) + c_{cov}\,p_c(t+1)\,p_c(t+1)^T$$

Rysunek z: *A. Auger, N. Hansen:*
CMA-ES — Evolution Strategies and Covariance Matrix Adaptation, tutorial, GECCO'2011

Modyfikacja macierzy kowariancji algorytm CMA-ES (wersja 2)

$C(1)=I$

Przestrzeń x

$while\,!\,stop$

$\quad generuj\ d_i(t)\sim N(0,C(t)),\, i=1\ldots\lambda$

$\quad oblicz\ \boxed{q_i}(t)=q\left(m(t)+\sigma(t)\cdot d_i(t)\right)$

$\quad sortuj\ według\ q_i(t)$

$\quad \Delta(t)=\frac{1}{\mu}\sum_{i=1}^{\mu} d_i(t)$

$\quad \boxed{m}(t+1)=m(t)+\sigma(t)\cdot\Delta(t)$

uwzględnienie wielu wektorów różnic

$\quad \boxed{C}(t+1)=(1-c_{cov})C(t)+c_{cov}\sum_{i=1}^{\mu} d_i(t)\,d_i(t)^T$

$\quad t\leftarrow t+1$

Wszystkie operacje w przestrzeni x

Pełny CMA-ES

$C(1)=I, \quad p_c(1)=0, \quad p_\sigma(1)=0$

$\mathit{while}\ !\ \mathit{stop}$

$\quad \mathit{generuj}\ \boxed{d_i(t)} \sim N(0, C(t)), i=1\ldots\lambda$

$\quad \mathit{oblicz}\ \boxed{q_i(t)=q(m(t)+\sigma(t)\cdot d_i(t))}$

$\quad \mathit{sortuj\ według}\ q_i(t)$

$$\boxed{\Delta}(t)=\frac{1}{\mu}\sum_{i=1}^{\mu} d_i(t)$$

$$\boxed{m}(t+1)=m(t)+\sigma(t)\cdot\Delta(t)$$

$$\boxed{p_\sigma}(t+1)=(1-c_\sigma)\,p_\sigma(t)+C^{-1/2}\sqrt{1-(1-c_\sigma)^2}\,\sqrt{\mu}\,\Delta(t)$$

$$\boxed{p_c}(t+1)=(1-c_c)\,p_c(t)+\sqrt{1-(1-c_c)^2}\,\sqrt{\mu}\,\Delta(t)$$

$$\boxed{\sigma}(t+1)=\sigma(t)\cdot\exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|p_\sigma\|}{E\|N(0,I)\|}-1\right)\right)$$

$$\boxed{C}(t+1)=(1-c_1-c_\mu)\,C(t)+c_1\,p_c(t+1)\,p_c(t+1)^T+c_\mu\sum_{i=1}^{\mu} d_i(t)\,d_i(t)^T$$

$t \leftarrow t+1$

Przestrzeń x

Przestrzeń skojarzona

Pełny CMA-ES

$$C(1)=I, \quad p_c(1)=0, \quad p_\sigma(1)=0$$

$$\textit{while}\,!\,\textit{stop}$$

$$\quad \textit{generuj}\; d_i(t) \sim N(0, C(t)), i=1\ldots\lambda$$

$$\quad \textit{oblicz}\; q_i(t)=q\left(m(t)+\sigma(t)\cdot d_i(t)\right)$$

$$\quad \textit{sortuj według}\; q_i(t)$$

$$\quad \Delta(t)=\frac{1}{\mu}\sum_{i=1}^{\mu} d_i(t)$$

$$\quad m(t+1)=m(t)+\sigma(t)\cdot\Delta(t)$$

$$\quad p_\sigma(t+1)=(1-c_\sigma)\,p_\sigma(t)+\sqrt{1-(1-c_\sigma)^2}\,\sqrt{\mu}\,C^{-1/2}\,\Delta(t)$$

$$\quad \sigma(t+1)=\sigma(t)\cdot\exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|p_\sigma\|}{E\|N(0,I)\|}-1\right)\right)$$

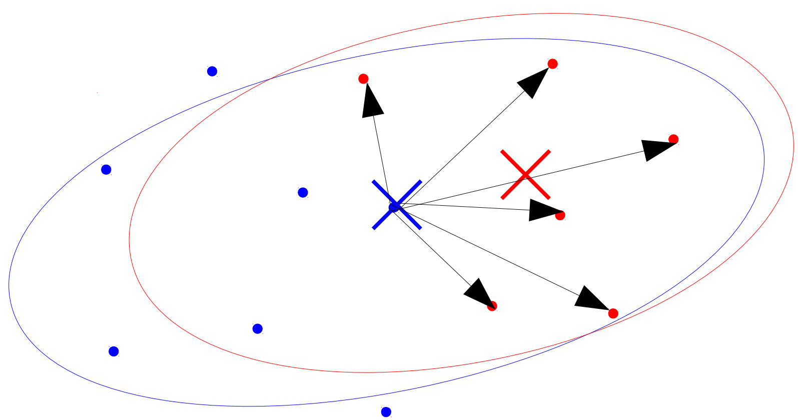
$$\quad p_c(t+1)=(1-c_c)\,p_c(t)+\sqrt{1-(1-c_c)^2}\,\sqrt{\mu}\,\Delta(t)$$

$$\quad C(t+1)=(1-c_1-c_\mu)\,C(t)+c_1\,p_c(t+1)\,p_c(t+1)^T+c_\mu\sum_{i=1}^{\mu} d_i(t)\,d_i(t)^T$$

$$\quad t \leftarrow t+1$$

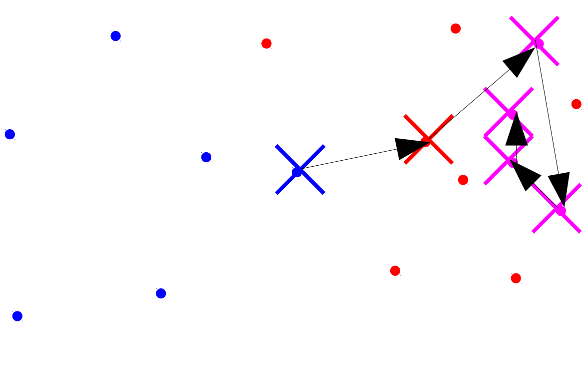
Beyer, Sendhoff, Simplify Your Covariance Matrix AdaptationEvolution Strategy, IEEE Trans on Evol. Comp. 2017

CMAES
Covariance Matrix Adaptation
Evolution Strategy

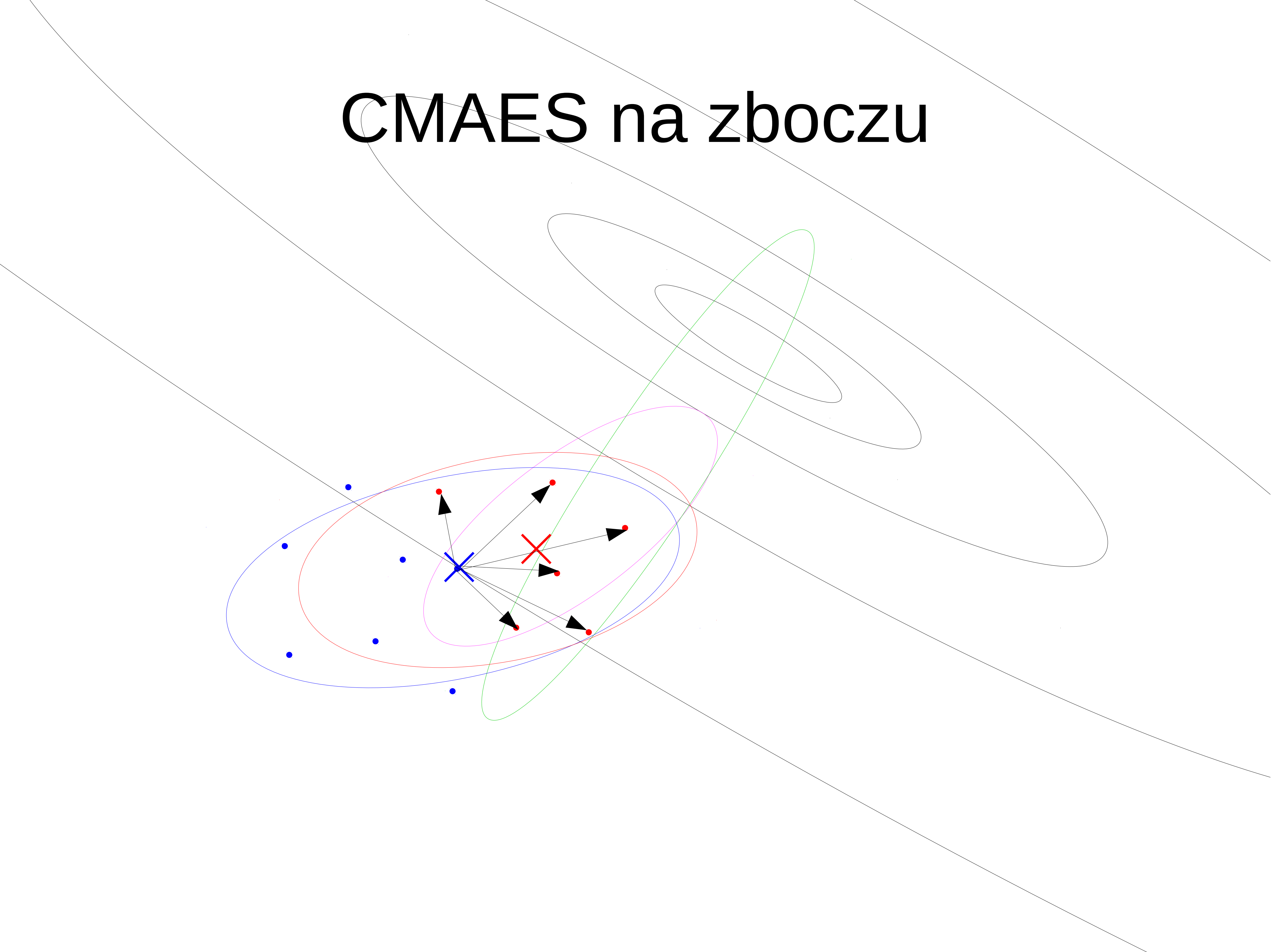


Na podstawie selekcji adaptuje się kształt macierzy kowariancji

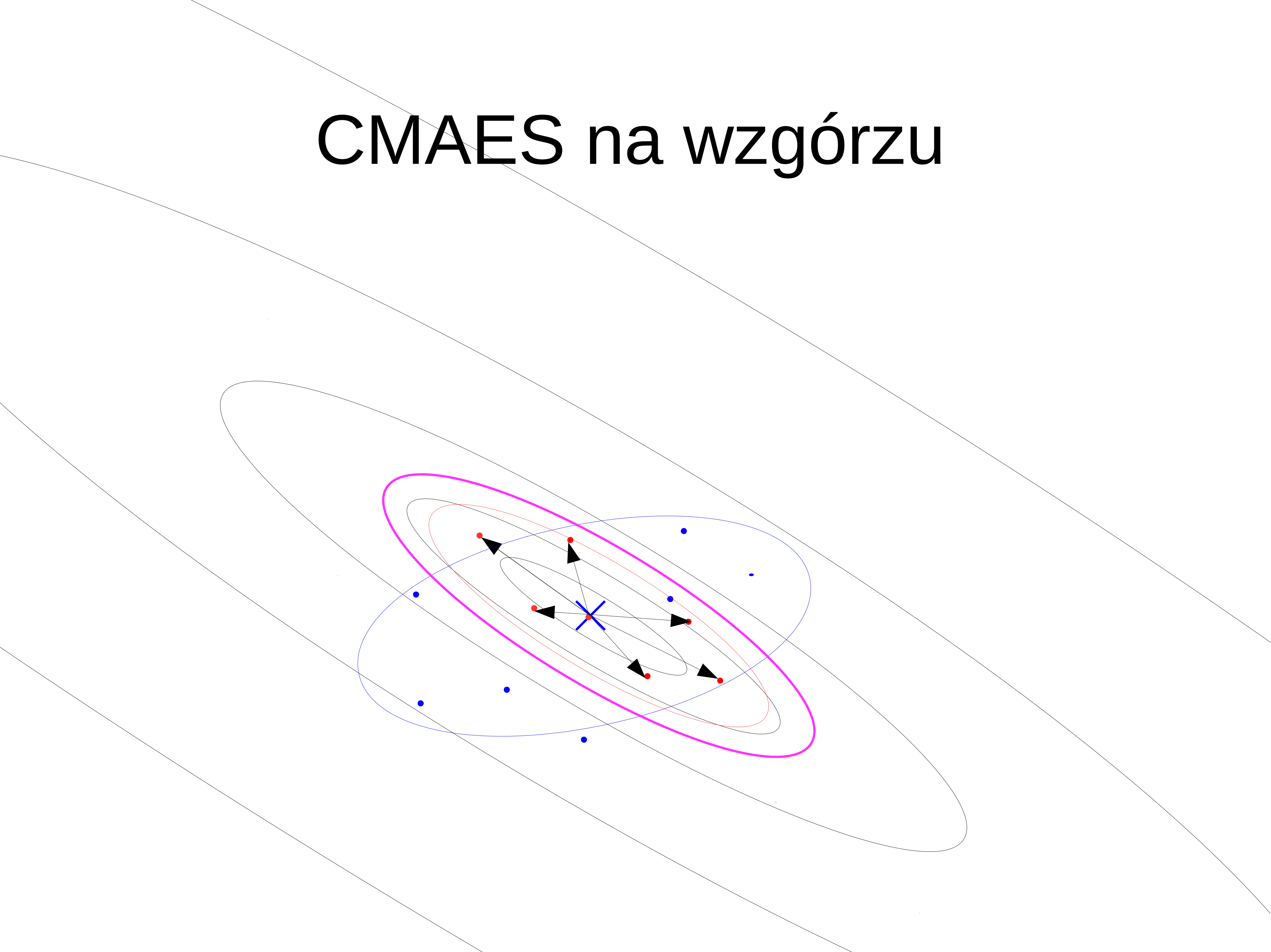
Jej skala zależy od ścieżki ewolucji

+  =CMAES

CMAES na zboczu



CMAES na wzgórzu



CMAES na wzgórzu