

**Computazione evolutivistica: algoritmi genetici e strategie evolutive - Fondamenti ed esempi di applicazione nel campo della Computer-Vision, Biomeccanica ed Elaborazione di Immagini Biomediche**

Cerveri Pietro

Dipartimento di Bioingegneria – Politecnico di Milano

25,27 Ottobre 2005

1/49

**Sommario**

**Le strategie evolutive di ottimizzazione**

Ampiezza della mutazione

Adattatività della mutazione

Determinazione dei parametri di mutazione

Esempi

2/49

## Strategie evolutive

- Sviluppate in Germania ~1970
- I. Rechenberg, H.-P. Schwefel
- Tipicamente applicata a:
  - Ottimizzazione numerica
- Caratteristiche principali:
  - veloce
  - buon risolutore per ottimizzazione a valori reali
  - Teoria base consolidata. Continui sviluppi teorici
- Potenzialità:
  - Autoadattamento della funzione di mutazione alla topologia della funzione obbiettivo

Trovo la soluzione ottimale, analizzando le soluzioni che via via vengono testate. Non vengono eseguite esplicitamente operazioni sulla funzione obbiettivo (e.g. calcolo di gradienti...).

3/49

## Caratteristiche tecniche

Rappresentazione cromosoma	Gli elementi della popolazione (padri e figli) sono vettori di numeri reali (le soluzioni)
Ricombinazione	Discrete o intermedia
Mutazione	Rumore Gaussiano
Selezione genitori	Casuale: distr. uniforme
Selezione candidati	$(\mu, \lambda)$ or $(\mu + \lambda)$
Auto-adattamento	Passo di mutazione

4/49

## Idee di base

Sono algoritmi iterativi.

Ad ogni generazione (iterazione) viene valutata una popolazione (insieme di soluzioni al problema) in base alla fitness degli elementi della popolazione esaminati (funzione obiettivo calcolata in ciascuna delle soluzioni esaminate).

All'iterazione successiva viene prodotta una nuova generazione, a partire dagli elementi della generazione precedente.

Vengono inseriti criteri robusti di ricerca che vanno a modificare, generazione per generazione, i parametri in base ai quali vengono generati i figli.

5/49

## Gli elementi della popolazione

- I cromosomi (o elementi della popolazione o genotipi o soluzioni candidate) sono costituiti da:
  - Variabili oggetto:  $x_1, \dots, x_n$
  - Variabili di strategia:
    - Passo di mutazione
    - Tasso di ricombinazione
    - Parametri di selezione
    - .....
- variabili oggetto  $x_i$  sulle quali verranno applicati gli operatori di ricombinazione mutazione e selezione fino a che non viene raggiunta una soluzione ottima del problema (la soluzione). Un individuo nella popolazione consiste in un genotipo che rappresenta un punto nello spazio di ricerca (cioè lo spazio delle possibili soluzioni).
- variabili di strategia  $S_i$  che determinano la dinamica dell'evoluzione della soluzione (di  $x_i$  - agiscono sugli operatori di ricombinazione, mutazione, e selezione)

6/49

## I genitori della prima generazione

- I genitori sono scelti da una distribuzione statistica, solitamente uniforme
- La selezione genitori in ES non è polarizzata – ciascun individuo ha la stessa probabilità di essere scelto
- Si noti che in ES genitore indica un membro di popolazione (in GA: indica un membro della popolazione selezionato per essere sottoposto a variazione), questo perchè potenzialmente tutti i membri della popolazione possono diventare genitori.
- Tra una generazione e la successiva avvengono gli eventi evolutivi:
  - Ricombinazione
  - Mutazione
  - Selezione

7/49

## Principi di ricombinazione

### Quali sono i genitori che diventano genitori nella generazione successiva?

A seconda che le soluzioni (gli elementi della popolazione) siano enumerati o possano assumere valori continui posso selezionare l'elemento intermedio o calcolare il valore medio (la posizione media).

### Selezione dei genitori per la produzione dei figli:

1. Estrae ad ogni iterazione (generazione) uno o più genitori.
2. Agisce sui genitori fissati
  - Mediando i valori dei geni corrispondenti ai genitori (nel caso comune di più genitori)
  - Selezionando uno dei geni tra i genitori

8/49

## Esempio di ricombinazione e tassonomia

Dati 2 genitori:  $x_i$  ed  $y_i$ , viene generato il figlio  $z_i$ .

		2 Genitori fissati	2 genitori selezionati per ogni $i$
Media	$z_i = (x_i + y_i)/2$	Intermedia locale	Intermedia globale
Selezione	$z_i$ is $x_i$ or $y_i$ scelta random	Discreta locale	Discreta globale

9/49

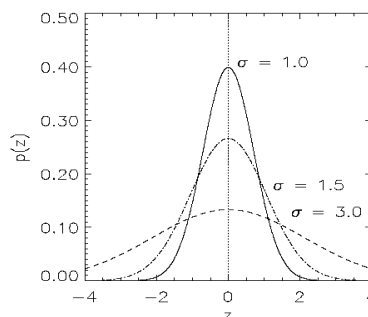
## Mutazione

$$\text{Figlio}^t = \text{Genitore}^t + z$$

$z$  – variabile stocastica che introduce variabilità nella specie (perturba la soluzione).

### Esempio:

- I valori  $z$  sono estratti dalla distribuzione normale:  $N(\xi, \sigma)$ 
  - la media  $\xi$  is set to 0
  - la deviazione standard  $\sigma$  è chiamata passo di mutazione



10/49

## Selezione

### Strategia Plus ( $\mu+\lambda$ ) e Strategia Comma ( $\mu,\lambda$ ).

• Nel primo caso anche i genitori possono partecipare alla selezione del genitore della generazione successiva, mentre nel secondo solo i figli possono essere selezionati, mentre i genitori muoiono.  $\mu$  rappresenta il numero di individui nella popolazione, mentre  $\lambda$  il numero di figli concepiti per ogni generazione.

Nella strategia **Plus** i migliori  $\mu$  individui tra padri e figli ( $\mu+\lambda$ ) sopravvivranno e diventeranno genitori nella generazione successiva, mentre nella **Comma** la selezione avviene solo tra i  $\lambda$  figli. In genere,  $\lambda > \mu$  (pressione della selezione,  $\lambda \approx 7 \cdot \mu$  è un setting usual).

Esclude in modo deterministico le soluzioni peggiori.

11/49

## Selezione (II)

- La selezione ( $\mu+\lambda$ ) è una strategia elitista
- La selezione ( $\mu,\lambda$ ) può “dimenticare”
- Spesso si preferisce la selezione ( $\mu,\lambda$ ) poichè:
  - Miglior comportamento nell’evitare ottimi locali
    - La strategia + risulta più sensibile a insiemi di genitori che hanno colonizzato ottimi locali
  - Miglior comportamento nel inseguire ottimi in movimento (processi non stazionari)
  - Integrata con matrice di covarianza produce su scala temporale media effetti quasi-elitisti

12/49

## Esempio I – (1+1) ES

- Obiettivo: minimizzare  $f(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}$
- Algoritmo: “two-membered ES” usando:
  - Vettori in  $\mathbb{R}^n$  direttamente come cromosoma
  - Dimensione della popolazione: 1
  - Non c’è ricombinazione.
  - Opera solo la mutazione che crea un singolo figlio
  - Valutazione della funzione per il padre e il figlio
  - Scelgo come padre della generazione successiva chi tra i due ha espresso una fitness migliore.

13/49

## Pseudo codice

- Nella prima generazione,  $t = 0$ , crea l’individuo iniziale  
 $\mathbf{x}^{t=0} = [x_1^{t=0}, \dots, x_n^{t=0}]$
- Ripeti fino a che la condizione di uscita non è verificata
  - Estrai un variabile  $z_i$  da una distribuzione normale (per tutti gli elementi del genotipo,  $i = 1, \dots, n$ ):
  - $y_i^t = x_i^t + z_i$
  - IF  $f(\mathbf{x}^t) < f(\mathbf{y}^t)$  THEN  $\mathbf{x}^{t+1} = \mathbf{x}^t$ 
    - ELSE  $\mathbf{x}^{t+1} = \mathbf{y}^t$
  - ENDIF
  - Passa alla generazione successiva:  $t = t+1$

14/49

## Sommario

Le strategie evolutive di ottimizzazione

Ampiezza della mutazione

Adattatività della mutazione

Determinazione dei parametri di mutazione

Esempi

15/49

## Adattamento della mutazione

- I valori  $z$  sono estratti dalla distribuzione  $N(\xi, \sigma)$ 
  - la media  $\xi$  is set to 0
  - la deviazione standard  $\sigma$  è chiamata passo di mutazione e definisce l'ampiezza dello spazio delle soluzioni all'interno del quale verranno generate le soluzioni esaminate nella generazione successiva.
- $\sigma$  viene aggiornata run-time con la regola euristica “1/5 success rule”:

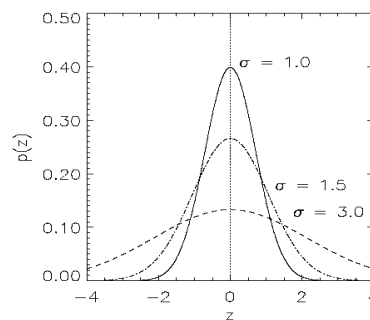
• Questa regola ricalcola  $\sigma$  ogni  $k$  epoche tramite:

$$-\sigma = \sigma / c \quad \text{se } p_s > 1/5$$

$$-\sigma = \sigma \cdot c \quad \text{se } p_s < 1/5$$

$$-\sigma = \sigma \quad \text{se } p_s = 1/5$$

dove  $p_s$  la percentuale di mutazioni con successo,  $0.8 \leq c \leq 1$



16/49



## Esempio di mutazione

OK = fitness del figlio migliore di quella di suo padre.

- Es:  $k=10$ 
  - $k = 1$  il figlio è stato selezionato mutazione OK
  - $k = 2$  il figlio è stato selezionato mutazione OK
  - $k = 3$  il figlio è stato selezionato mutazione nOK
  - $k = 4$  il figlio non è stato selezionato mutazione OK
  - $k = 5$  il figlio è stato selezionato mutazione nOK
  - $k = 6$  il figlio è stato selezionato mutazione OK
  - $k = 7$  il figlio non è stato selezionato mutazione nOK
  - $k = 8$  il figlio è stato selezionato mutazione nOK
  - $k = 9$  il figlio è stato selezionato mutazione OK
  - $k = 10$  il figlio è stato selezionato mutazione OK
- $0.60 > 0.2$  quindi  $\sigma = \sigma / c$

17/49

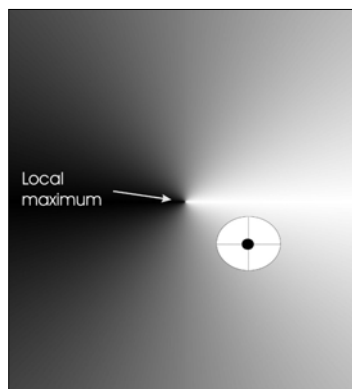
## Mutazione – $\sigma$ singolo

- Ho definito un unico valore di  $\sigma$ , comune a tutte le variabili che compongono il vettore degli elementi della popolazione.
- Ogni variabile costituente l'elemento viene mutata in modo scorrelato alle altre.
- Mutazione scorrelata con singolo valore  $\sigma$ 
  - Cromosomi:  $[x_1, \dots, x_n; \sigma]$
  - $\sigma' = \sigma \cdot \exp(\tau \cdot N(0,1))$
  - $x'_i = x_i + \sigma' \cdot N(0,1)$
  - Tipicamente il learning rate  $\tau \propto 1/n^{1/2}$   $\tau < 0$
  - Limiti:  $\sigma' < \varepsilon_0 \Rightarrow \sigma' = \varepsilon_0$  (per evitare che la regione di ricerca della soluzione collassi in un punto).

Ricordiamo che ciascun elemento della popolazione (cromosoma) è caratterizzato dalle variabili oggetto (soluzioni) + variabili strategia )parametri di evoluzione che sono associabili agli operatori di ricombinazione mutazione e selezione).

18/49

I figli generati per mutazione sono statisticamente distribuiti nello spazio delle soluzioni all'interno di un'ipersfera (distribuzione isotropica)



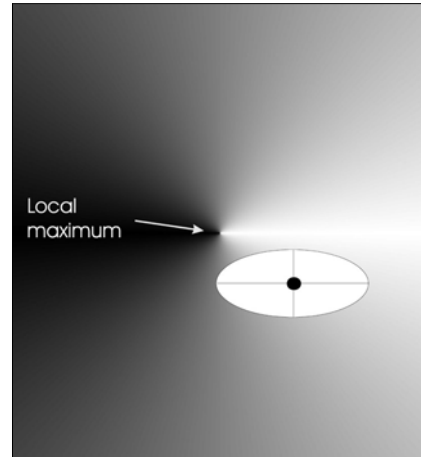
19/49

## Mutazione – $\sigma$ personalizzato

- Inserisco valori diversi di  $\sigma$  per diverse variabili. Concedo che lo spazio delle soluzioni sia esplorato maggiormente lungo alcune direzioni rispetto altre.
- Mutazione scorrelata con molteplici valori  $\sigma_i$ 
  - Cromosomi:  $[x_1, \dots, x_n, \sigma_1, \dots, \sigma_n]$
  - $\sigma'_i = \sigma_i \cdot \exp[\tau' \cdot N(0,1) + \tau_i \cdot N_i(0,1)]$
  - $x'_i = x_i + \sigma'_i \cdot N_i(0,1)$
  - $\mathbf{x}' = \mathbf{x} + \exp[\tau' \cdot N(0, \mathbf{D})]$
  - 2 parametri di learning:
    - $\tau'$  learning rate globale
    - $\tau_i$  learning rate relativo alla specifica variabile i-esima
  - Valori tipici:  $\tau_i \propto 1/(2n)^{1/2}$  and  $\tau' \propto 1/(2n^{1/2})^{1/2}$   $\tau_i < 0$   $\tau' < 0$  dove  $n$  è il numero di generazioni.
  - Limiti:  $\sigma'_i < \epsilon_0 \Rightarrow \sigma'_i = \epsilon_0$   
*Favorisco le variazioni degli elementi della popolazione lungo le direzioni dove  $\sigma_i$  è maggiore.*  
 *$\sigma$  decresce esponenzialmente con il numero di generazioni.*

20/49

I figli generati per mutazione sono statisticamente distribuiti all'interno di un ellissoide con i raggi paralleli agli assi (alle variabili, agli elementi della popolazione) ed i cui raggi sono dati dai  $\sigma_i$ .



21/49

## Sommario

Le strategie evolutive di ottimizzazione

Ampiezza della mutazione

**Adattatività della mutazione**

Determinazione dei parametri di mutazione

Esempi

22/49

## Mutazione – $\sigma$ personalizzato

- Inserisco valori diversi di  $\sigma$  per diverse variabili. Concedo che lo spazio delle soluzioni sia esplorato maggiormente lungo alcune direzioni rispetto altre.

*Favorisco le variazioni degli elementi della popolazione lungo le direzioni dove  $\sigma_i$  è maggiore.*

*$\sigma$  decresce esponenzialmente con il numero di generazioni.*

*Vorrei trovare un modo di adattare l'ampiezza della regione di ricerca nelle diverse direzioni (variabili) in funzione della fitness (bontà) della soluzione per quanto riguarda quella direzione (variabile).*

23/49

## Richiamo di statistica

Data una variabile casuale,  $x$ , il suo valore medio calcolato su  $N$  campioni è dato da:

$$M_x = \frac{\sum_{k=1}^N x_k}{N}$$

Data una variabile casuale,  $x$ , la sua varianza su  $N$  campioni è data da:

$$\sigma^2_x = \frac{\sum_{k=1}^N (x_k - M_x)^2}{N}$$

Data una variabile casuale,  $x$ , la sua deviazione standard su  $N$  campioni è data da:

$$\sigma_x = \sqrt{\frac{\sum_{k=1}^N (x_k - M_x)^2}{N}}$$

*Varianza e deviazione standard descrivono la dispersione attorno al valor medio.*

24/49

## Correlazione

Date due variabili casuali:  $x_i, x_j$ , l'indice di correlazione misura quanto le coppie di variabili estratte:  $p(x_i, x_j)$  stanno su una retta:

$$r = \frac{M_{x_i x_j} - M_{x_i} M_{x_j}}{\sigma_{x_i} \sigma_{x_j}} \quad -1 \leq r \leq +1$$

Definendo la covarianza tra  $x_i$  ed  $x_j$  come:

$$\sigma_{x_i x_j} = \frac{1}{N} \sum_i \sum_j (x_i - M_{x_i})(x_j - M_{x_j})$$

Dalla definizione di deviazione standard risulta:

$$r = \frac{\sigma_{x_i x_j}}{\sigma_{x_i} \sigma_{x_j}}$$

25/49

## Correlazione tra più variabili

Date N variabili casuali:  $x = [x_1, x_2, \dots, x_N]$  si può misurare la correlazione tra coppie di variabili. E' comodo rappresentare la correlazione in un'unica matrice detta matrice di covarianza come:

$$C = \begin{bmatrix} \sigma_{x_1 x_1} & \sigma_{x_1 x_2} & \cdot & \sigma_{x_1 x_N} \\ \sigma_{x_2 x_1} & \sigma_{x_2 x_2} & \cdot & \sigma_{x_2 x_N} \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{x_N x_1} & \sigma_{x_N x_2} & \cdot & \sigma_{x_N x_N} \end{bmatrix}$$

Varianza:  $\sigma_{x_i x_i} = \sigma_{x_i}^2$                       N parametri

Covarianza:  $\sigma_{x_i x_j} = \sigma_{x_j x_i} \quad i \neq j$                       (N-1)<sup>2</sup>/2 parametri

26/49

## Diagonalizzazione di una matrice

Una matrice quadrata i cui autovalori siano distinti è diagonalizzabile. La matrice di Covarianza ricade sotto queste ipotesi.

$$C = B W B^T \quad \rightarrow \quad C = B D^2 B^T$$

**B** è matrice ortonormale, **W** è matrice diagonale e contiene gli autovalori di **C**.

### Significato geometrico

- **B** è una matrice di rotazione.
- Il sistema di riferimento viene ruotato in modo tale che i suoi assi siano orientati secondo le direzioni principali di **C**, ovverosia le direzioni ortogonali lungo le quali si ha la massima variabilità.

27/49

## Elementi per la Mutazione – $\sigma$ correlato

- Si preferisce esprimere i valori off-diagonal della matrice di covarianza in funzione delle varianze in questo modo:
- Mutazioni correlate
  - Cromosomi:  $[x_1, \dots, x_n; \sigma_1, \dots, \sigma_n; \alpha_{12}, \alpha_{13}, \dots, \alpha_{N-1, N-1}]$ .
  - e la matrice di covarianza  $C_{n \times n}$  gode di queste proprietà:
    - positiva
    - $c_{ii} = \sigma_i^2$
    - $c_{ij} = 0$  se  $i$  e  $j$  sono scorrelati.
    - $c_{ij} \neq 0$  altrimenti, rappresentati dagli  $\alpha_{ij}^2$ .

28/49

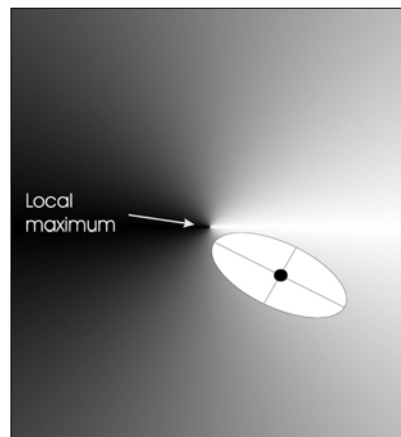
## Mutazione – $\sigma$ correlato

- Meccanismo di mutazione:
  - $\sigma'_i = \sigma_i \cdot \exp[\tau' \cdot N(0,1) + \tau_i \cdot N_i(0,1)]$
  - $\alpha'_{ij} = \alpha_{ij} + \beta \cdot N(0,1)$
  - $\mathbf{x}' = \mathbf{x} + N(\mathbf{0}, \mathbf{C}')$ 
    - $\mathbf{x}$  rappresenta il vettore  $\langle x_1, \dots, x_n \rangle$
    - $\mathbf{C}'$  è la matrice di covarianza  $\mathbf{C}$  dopo la mutazione dei valori  $\alpha$  e  $\sigma$ .
- Valori tipici:  $\tau_i \propto 1/(2n)^{1/2}$  e  $\tau' \propto 1/(2n^{1/2})^{1/2}$  and  $\beta \approx 5$
- Nota bene:
  - Il valore di  $\sigma$  determina l'ampiezza della regione di ricerca nelle diverse direzioni ( $\sigma_i$ ).
  - La regione ellissoidale viene poi ruotata utilizzando la matrice di covarianza.

29/49

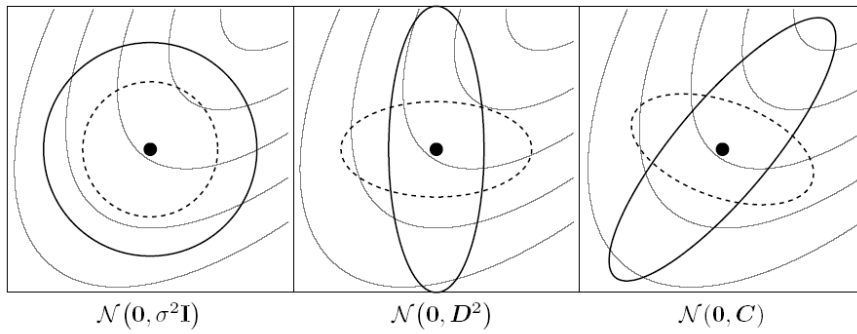
I figli generati per mutazione sono statisticamente distribuiti all'interno dell'iperellissoide orientato lungo le direzioni principali della varianza.

Questo consente di adattare la regione di ricerca al meglio alla topologia della funzione obiettivo senza doverla determinare in modo esplicito.



30/49

## Sintesi



31/49

## Sommario

Le strategie evolutive di ottimizzazione

Ampiezza della mutazione

Adattatività della mutazione

**Determinazione dei parametri di mutazione**

Esempi

32/49



## Ricombinazione

Scelta dei genitori per la generazione successiva:  $m^{(g)}$ .

- $m^{(g)}$  viene calcolato come media (pesata) di  $L$  individui selezionati tra i figli:  $x_1^{(g)}, \dots, x_\lambda^{(g)}$

$$\mathbf{m}^{(g)} = \sum_{i=1}^L w_i \mathbf{x}_i^{(g)}$$

$$\sum_{i=1}^L w_i = 1$$

- $\mu \leq \lambda$  è la dimensione della popolazione
- $w_i$  coefficienti  $> 0$
- $x_i^{(g)}$   $i$ -esimo migliore individuo tra  $x_1^{(g)}, \dots, x_\lambda^{(g)}$ 
  - $f(x_1^{(g)}) \leq f(x_2^{(g)}) \dots \leq f(x_\lambda^{(g)})$
  - $f$  è la funzione da minimizzare: fitness.
- Tipicamente  $w_i \propto L - i + 1$  e  $L \approx \lambda/2$

33/49

## Mutazione

$$\mathbf{x}_k^{(g+1)} = \mathbf{m}^{(g)} + \rho^{(g)} \mathbf{B} \mathbf{D} \mathbf{z}_k$$

$$\mathbf{z}_k = N(\mathbf{0}, \mathbf{I})$$

$$\mathbf{z}_k = [z_1 = N(0,1), z_2 = N(0,1), \dots, z_n = N(0,1)]^T$$

- Vengono generati  $\lambda$  figli
- $\mathbf{x}^{(g+1)}$  sarà funzione dei padri selezionati
- $\rho$  è globale: global step size

*Devo determinare:  $C$  e  $\rho$ . Per questo utilizzo l' "evolution path"*

34/49

## Derandomizzazione tramite Evolution Path

- Sia la matrice di covarianza  $\mathbf{C}$  che il global step size  $\sigma$  sono aggiornati indirettamente tramite il cosiddetto evolution path  $\mathbf{s}$ .

$$\mathbf{s}^{(g+1)} = (1 - a_s) \mathbf{s}^{(g)} + a_u \mathbf{B}^{(g)} \mathbf{G}^{(g)} \mathbf{z}_{i_{sel}}^{(g)} \quad \mathbf{z}_{i_{sel}} \text{ rappresenta il genitore migliore}$$

$$a_u^2 + (1 - a_s)^2 = 1$$

$$\text{Nuovo\_valore} = a * \text{Vecchio\_valore} + b * \text{Innovazione}$$

Cf. Discounted reward in RL

*s unisce tutti i genitori migliori. Dà un'idea dell'evoluzione. Path evolutivo all'interno dello spazio delle soluzioni. Connette i genitori migliori.*

35/49

## Aggiornamento della matrice di Covarianza

$$\text{Nuovo\_valore} = a * \text{Vecchio\_valore} + b * \text{Innovazione}$$

$$\mathbf{C}^{(g+1)} = (1 - \alpha_C) \mathbf{C}^{(g)} + \alpha_C \mathbf{S}^{(g+1)}$$

$$\mathbf{S}^{(g+1)} = \mathbf{s} \mathbf{s}^T$$

Calcolato sul path evolutivo

*Cosa vuol rappresentare questa equazione di aggiornamento?*

36/49

## Adattamento del global step size $\rho$

$\sigma_\rho^{(g+1)}$  • Evolution path normalizzato (contiene solamente informazioni sulla direzione).

$$\sigma_\rho^{(g+1)} = (1 - \alpha_s) \mathbf{s}_\rho^{(g)} + \alpha_s \mathbf{B}^{(g)} \mathbf{z}_{i_{sel}}^{(g)}$$

$\bar{e}_n = \left( 1 - \frac{1}{4n} + \frac{1}{21n^2} \right)$  • Approssimazione al 2° ordine del valore atteso della distribuzione della lunghezza dei vettori estratti da  $N(\mathbf{0}, \mathbf{I})$

$\rho^{(g+1)} = \rho^{(g)} e^{\gamma \left( \|\mathbf{s}_\rho^{(g+1)}\| - \bar{e} \right)}$  □  $\rho$  decresce quando la direzione dell'evolution path cambia frequentemente. Quando una stessa direzione è ripetutamente selezionata  $\rho$  è incrementato.

$\gamma = 0.1$

*Se l'evoluzione cambia spesso direzione, meglio generare figli più vicini (e saranno sparsi lungo tutte le direzioni).*

37/49

## I parametri secondari

$$\mathbf{x}_k^{(g+1)} = \mathbf{m}^{(g)} + \rho^{(g)} \mathbf{B} \mathbf{D} \mathbf{z}_k$$

$$\mathbf{s}^{(g+1)} = (1 - a_s) \mathbf{s}^{(g)} + a_u \mathbf{B}^{(g)} \mathbf{G}^{(g)} \mathbf{z}_{i_{sel}}^{(g)}$$

$$\mathbf{s}_\rho^{(g+1)} = (1 - \alpha_s) \mathbf{s}_\rho^{(g)} + \alpha_s \mathbf{B}^{(g)} \mathbf{z}_{i_{sel}}^{(g)}$$

$$a_u^2 + (1 - a_s)^2 = 1$$

$$\mathbf{C}^{(g+1)} = (1 - \alpha_C) \mathbf{C}^{(g)} + \alpha_C \mathbf{S}^{(g+1)}$$

$$\mathbf{S}^{(g+1)} = \mathbf{s} \mathbf{s}^T$$

$$\rho^{(g+1)} = \rho^{(g)} e^{\gamma \left( \|\mathbf{s}_\rho^{(g+1)}\| - \bar{e} \right)}$$

*Devo determinare:  $\alpha_C, \alpha_s, \alpha_u$*

38/49

## Aggiornamento di $\alpha_s$

Regola l'aggiornamento di  $\rho$ . Questo deve essere particolarmente sensibile all'innovazione, in modo da pilotare le generazioni future dove più serve.

$$\alpha_s = 1/\sqrt{n}$$

Dove  $n$  è la dimensionalità dello spazio della popolazione

$$\alpha_C = 1/n^2$$

Il numero di parametri liberi della matrice di covarianza è dell'ordine di  $n^2 \rightarrow 1/n^2$ .

Rappresentano due fenomeni diversi:  $r$  è sensibile alle variazioni rapide della fitness della popolazione, mentre  $C$  è più sensibile alle variazioni lente, filtrando perciò le variazioni brusche occasionali.

- Es:  $n=10$ 
  - $c_{cov}=0.01, \alpha_u=0.72$
- La matrice di covarianza ha un *time span* molto più ampio dell'evolution path  $O(n^2)$
- L'evolution path può subire cambiamenti veloci mentre la matrice di covarianza ha bisogno di molto più tempo per fissare la memoria della direzione (deve essere selezionata molte volte)

39/49

selezione all'epoca  $g$

Calcola evolution path normalizzato  
 $s_p$

Aggiorna evolution path normalizzato

Aggiorna global step size  $\rho$

Aggiorna evolution path  $s$

Calcola matrice di covarianza corrente  
 $S=ss^T$

Aggiorna matrice di covarianza  $C$

**Schema di flusso dell'adattamento della mutazione**

40/49

## Sommario

Le strategie evolutive di ottimizzazione

Ampiezza della mutazione

Adattatività della mutazione

Determinazione dei parametri di mutazione

Esempi

41/49

### Problema 1: Calibrazione geometrica di una coppia stereo di telecamere

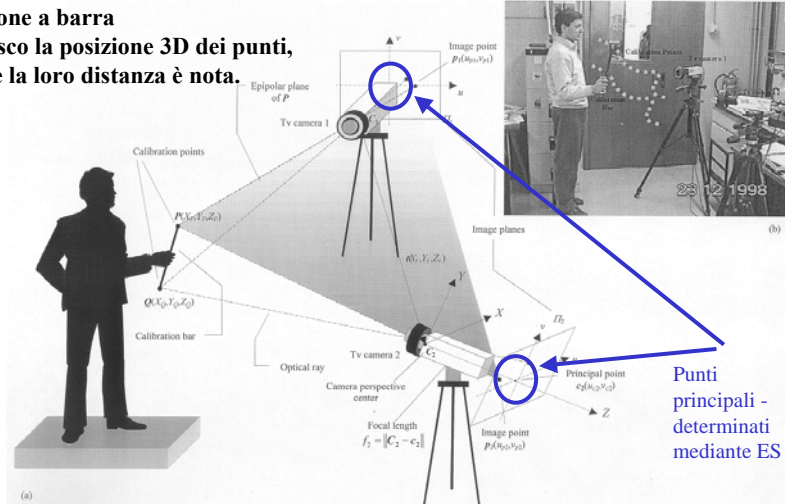
- Determinazione dei parametri geometrici del set-up:
  - Posizione e orientamento relativi di una coppia di camera.
  - Lunghezza focale e punto principale di ciascuna camera.
- Utilizzando la geometria epipolare
  - Relazione tra punti omologhi sulle due telecamere
  - Matrice fondamentale 3x3 (7 parametri liberi)
    - 3 rotazioni, 3 traslazioni a meno di un fattore di scala (2Dofs), 2 lunghezze focali
  - Consente di calcolare 7 parametri in forma chiusa utilizzando la geometria proiettiva.
- Punti principali
  - Stimo a posteriori i 2 punti principali,  $\mathbf{c}_1$  e  $\mathbf{c}_2$ , tramite Ricerca evolutiva

Spazio della ricerca quadridimensionale,  $\mathbf{c}_1[x_1, y_1]$  e  $\mathbf{c}_2[x_2, y_2]$  (“mondo” della popolazione).

42/49

# Il set-up

Calibrazione a barra  
 Non conosco la posizione 3D dei punti,  
 solamente la loro distanza è nota.

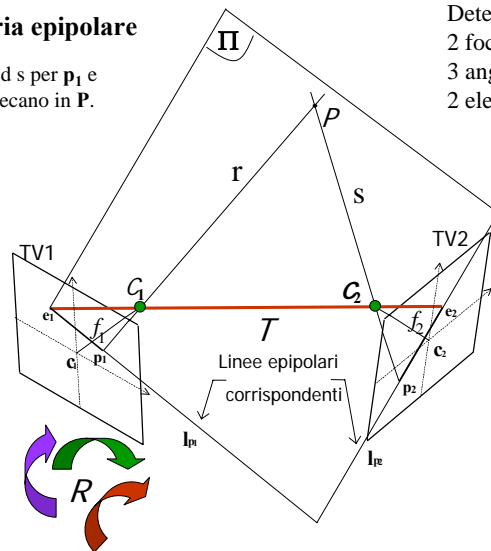


Punti principali - determinati mediante ES

## Geometria epolare

Le rette  $r$  ed  $s$  per  $p_1$  e  $p_2$  si intersecano in  $P$ .

Determino 7 parametri:  
 2 focali  
 3 angoli di orientamento  
 2 elementi del vettore posizione relativa



I parametri sono contenuti nella matrice  $F$

$$\overline{PC_2} \cdot \overline{C_1 C_2} \wedge \overline{PC_1} = p_2^T K_2^{-T} R(T \wedge) K_1^{-1} p_1 = 0 \Rightarrow p_2^T F p_1 = 0$$

$p_1$  e  $p_2$  sono misurati rispetto a  $c_1$  e  $c_2$  non noti.

## Soluzione tramite “Evolution strategies”

Procedura: ES(1,10)

1. scelgo una possibile soluzione per la coppia di punti principali
2. calcolo 7 parametri geometrici in forma chiusa, tramite matrice fondamentale.
3. effettuo la ricostruzione 3D e calcolo il fattore di scala (1 parametro).
4. Valuto la fitness della soluzione.

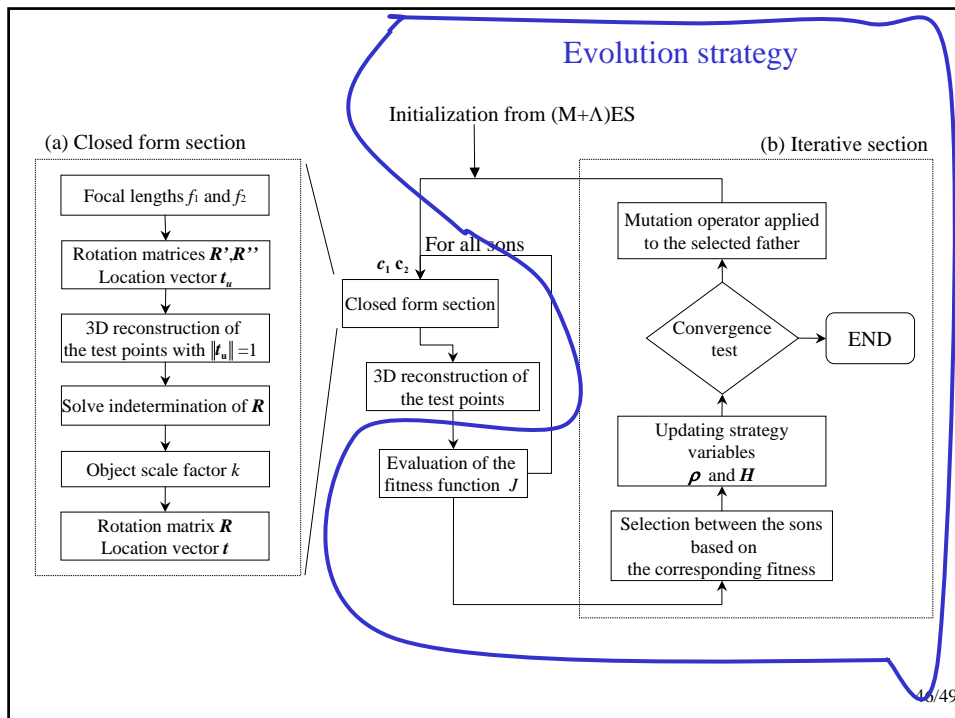
Funzione fitness: 
$$J = \alpha \sqrt{\frac{1}{N} \sum_{j=1}^N (d_j - L)^2} + \beta \sqrt{\frac{1}{N} \sum_{j=1}^N \left( \frac{e_{\text{int } j_r}^2 + e_{\text{int } j_o}^2}{2} \right)}$$

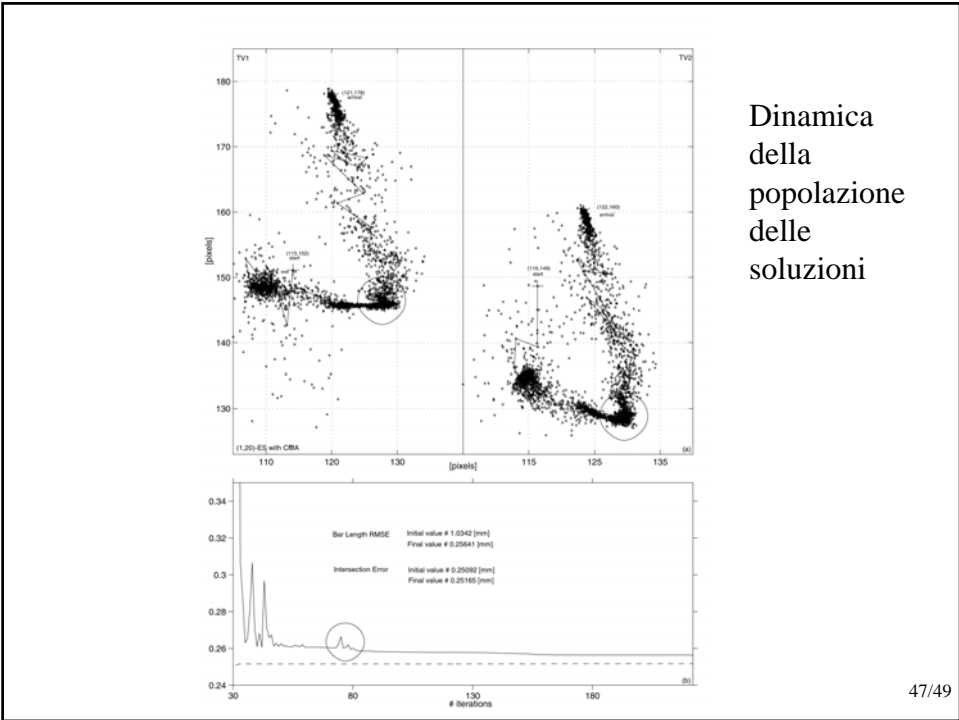
Errore sulla lunghezza della barra. Errore di intersezione (tra r ed s)

$$\min_{S_2=\{c_1, c_2\}} J(\{c_1, c_2\} | \{p_1\}, \{p_2\})$$

Elemento della popolazione:  $\left\{ \mathbf{w} \left[ \mathbf{c}_1(u_{c_1}, v_{c_1}), \mathbf{c}_2(u_{c_2}, v_{c_2}) \right]^T \right\}$

45/49





Dinamica  
della  
popolazione  
delle  
soluzioni

47/49

# Progetti

Come progetto per il corso potete implementare una variante di strategia evolutiva da applicare a problemi reali (e.g. routing, scheduling, ...).

48/49



## **Sommario**

Le strategie evolutive di ottimizzazione

Ampiezza della mutazione

Adattatività della mutazione

Determinazione dei parametri di mutazione

Esempi