

Derivative free optimization and evolutionary strategies

**Solution 1: Coordinate descent**

$$\begin{aligned}\mathcal{R}_{\text{emp}}(\boldsymbol{\theta}) &= \frac{1}{2} \|\mathbf{X}\boldsymbol{\theta} - \mathbf{y}\|_2^2 + \frac{1}{2} \|\boldsymbol{\theta}\|_2^2 = \frac{1}{2} \mathbf{y}^\top \mathbf{y} - \mathbf{y}^\top \mathbf{X}\boldsymbol{\theta} + \frac{1}{2} \boldsymbol{\theta}^\top \boldsymbol{\theta} + \frac{\lambda}{2} \boldsymbol{\theta}^\top \boldsymbol{\theta} \\ &= \frac{1}{2} \mathbf{y}^\top \mathbf{y} - \sum_{j=1}^p \mathbf{y}^\top \mathbf{X}_j \theta_j + \frac{1}{2} (1 + \lambda) \boldsymbol{\theta}^\top \boldsymbol{\theta} \\ \frac{\partial \mathcal{R}_{\text{emp}}}{\partial \theta_j} &= (1 + \lambda) \theta_j - \mathbf{y}^\top \mathbf{X}_j \stackrel{!}{=} 0 \\ \Rightarrow \theta_j^* &= \frac{\mathbf{y}^\top \mathbf{X}_j}{1 + \lambda}\end{aligned}$$

**Solution 2: CMA-ES**

Pick  $\mu = 3$  parents with highest fitness values, i.e.,  $\text{Id} = 1, 2, 5$  which we denote with  $\mathbf{x}_{1:\mu}$  and respective weights  $w_i = \frac{f_i}{\sum_{i=1}^{\mu} f_i} \approx (0.432, 0.265, 0.303)$ .

$$\begin{aligned}\mathbf{m}^{[1]} &= \mathbf{m}^{[0]} + 0.5 \sum_{i=1}^3 w_i (\mathbf{x}_i - \mathbf{m}^{[0]}) \approx (1.05, 0.84)^\top \\ \mathbf{C}_\mu &= \frac{1}{3-1} \sum_{i=1}^3 (\mathbf{x}_i - \mathbf{m}^{[0]})(\mathbf{x}_i - \mathbf{m}^{[0]})^\top \\ &\approx \begin{pmatrix} 0.187 & -0.617 \\ -0.617 & 2.139 \end{pmatrix} \\ \mathbf{C}^{[1]} &= 0.9 \cdot \mathbf{I}_3 + 0.1 \cdot \mathbf{C}_\mu \\ &\approx \begin{pmatrix} 0.919 & -0.062 \\ -0.062 & 1.114 \end{pmatrix}\end{aligned}$$