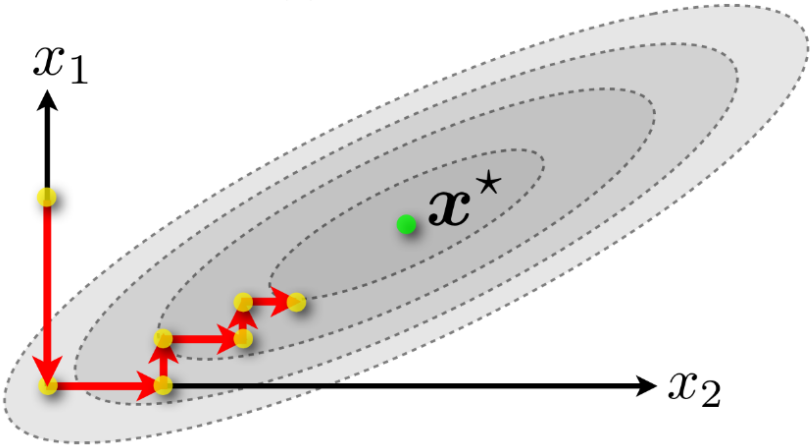
Coordinate descent

Axel Böhm

January 7, 2022

Coordinate Descent

Goal: Find $x^* \in \mathbb{R}^d$ minimizing $f(x)$.



Figure

Coordinate Descent

Modify only one coordinate per step:

$$\text{select } i_k \in \{1, \dots, d\}$$
$$x_{k+1} = x_k + \gamma e_{i_k}$$

where e_i is the i-th unit basis vector. Two main variants:

- ◇ Gradient-based stepsize:

$$x_{k+1} = x_k - \frac{1}{L} \nabla_{i_k} f(x_k) e_{i_k}$$

- ◇ Exact coordinate minimization:
  Solve the scalar problem $\arg\min_{\gamma \in \mathbb{R}} f(x_k + \gamma e_{i_k})$.
  - *hyperparameter free*

Randomized Coordinate Descent

select $i_k \in \{1, \dots, d\}$ uniformly at random

$$x_{k+1} = x_k + \gamma e_{i_k}$$

- Faster convergence than gradient descent
  (if coordinate step is d times cheaper than full gradient step)