University of Amsterdam
Informatics Institute

AMLAB
Amsterdam Machine Learning Lab

Machine Learning 1

Lecture 9.3 - Unsupervised Learning
Intermezzo: Lagrange Multipliers

Erik Bekkers

(Bishop Appendix E)

Slide credits: Patrick Forré and Rianne van den Berg

Image credit: Kirillm | Getty Images

Intermezzo: Lagrange Multipliers

- **In general:** Find maximum of $f(\mathbf{x})$ subject to $\overbrace{g(\mathbf{x})}^{\text{level set}} = c$ (2)

  e.g. $\|\mathbf{x}\| = c$

- Useful property: $\nabla g(\mathbf{x})$ is perpendicular to the constraint surface (App E)

- At constrained maximum, $\nabla f(\mathbf{x})$ must also be perpendicular to constraint surface

- (1): Therefore: $\nabla f(\boldsymbol{x}) + \lambda \nabla g(\boldsymbol{x}) = 0$

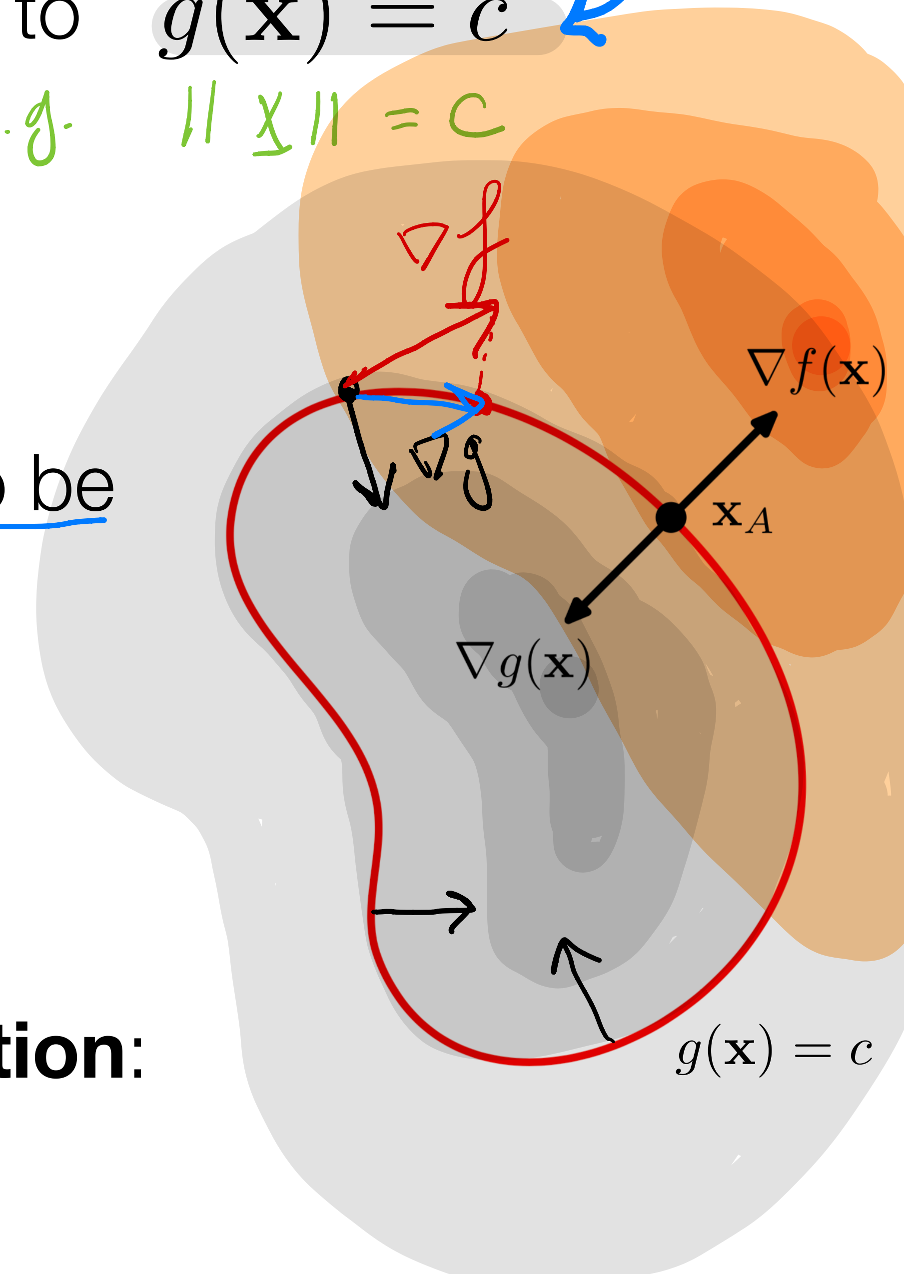
  λ : Lagrange multiplier

- It is helpful to introduce a **Lagrangian function**:

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda(g(\mathbf{x}) - c)$$

- Solutions to original problem: stationary points of $L(\mathbf{x}, \lambda)$

$$\frac{\partial}{\partial \boldsymbol{x}} L(\boldsymbol{x}, \lambda) = 0 \quad (1) \qquad \frac{\partial}{\partial \lambda} L(\boldsymbol{x}, \lambda) = 0 \quad (2)$$

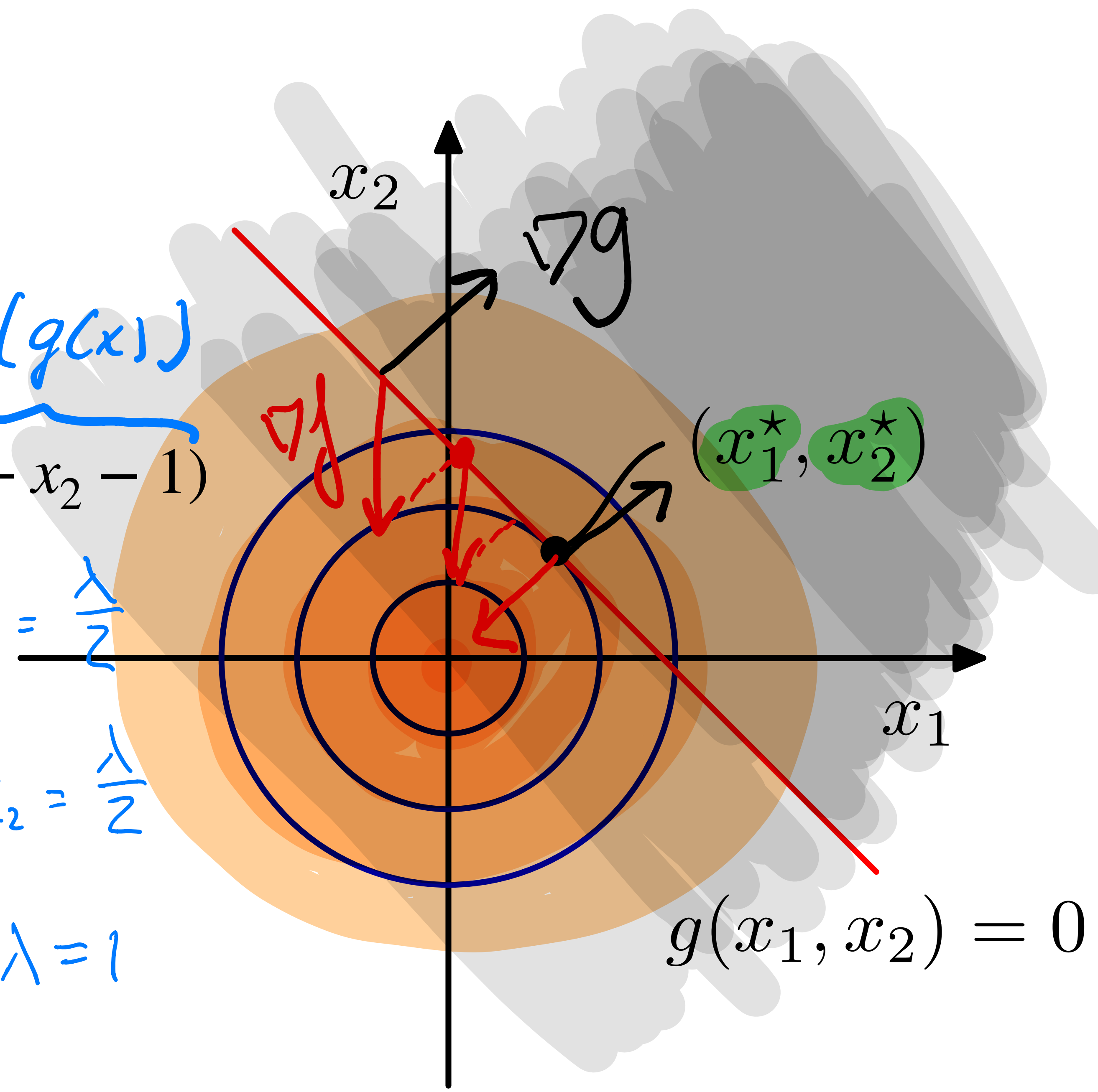


Lagrange multipliers: example

goal: $\max_{x_1, x_2} f(x_1, x_2)$ **s.t.** $g(x_1, x_2) = 0$

$$f(x_1, x_2) = 1 - x_1^2 - x_2^2$$

$$g(x_1, x_2) = x_1 + x_2 - 1$$



Lagrangian:

$$L(x_1, x_2, \lambda) = \overbrace{1 - x_1^2 - x_2^2}^{f(x_1, x_2)} + \overbrace{\lambda(x_1 + x_2 - 1)}^{\lambda(g(x))}$$

$$\frac{\partial}{\partial x_1} L(x_1, x_2, \lambda) = -2x_1 + \lambda = 0 : x_1 = \frac{\lambda}{2}$$

$$\frac{\partial}{\partial x_2} L(x_1, x_2, \lambda) = -2x_2 + \lambda = 0 : x_2 = \frac{\lambda}{2}$$

$$\frac{\partial}{\partial \lambda} L(x_1, x_2, \lambda) = x_1 + x_2 - 1 = 0 : \lambda = 1$$

$$\lambda = 1 \qquad x_1^* = \frac{1}{2} \qquad x_2^* = \frac{1}{2}$$