



Chương 4: Bài toán tối ưu

4.2 Thiết lập bài toán tối ưu

Nguyễn Văn Hời

Trường Đại học Công nghệ Thông tin
Bộ môn Toán - Lý





Nội dung

- Tổng quan.



Nội dung

- Tổng quan.
- Bài toán Joker.



Nội dung

- Tổng quan.
- Bài toán Joker.
- Bài toán quay rào.



Nội dung

- Tổng quan.
- Bài toán Joker.
- Bài toán quay rào.
- Bài toán container.



Nội dung

- Tổng quan.
- Bài toán Joker.
- Bài toán quay rào.
- Bài toán container.
- Mô hình hồi quy tuyến tính.



Nội dung

- Tổng quan.
- Bài toán Joker.
- Bài toán quay rào.
- Bài toán container.
- Mô hình hồi quy tuyến tính.
- Vec tơ hỗ trợ học máy.



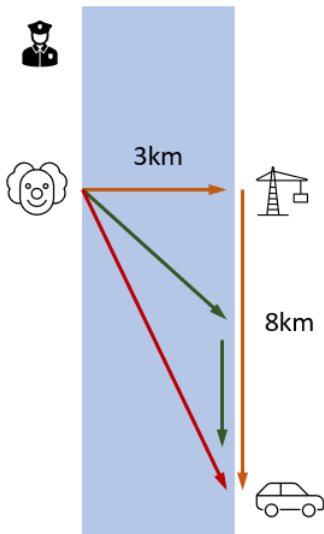
Nội dung

- Tổng quan.
- Bài toán Joker.
- Bài toán quay rào.
- Bài toán container.
- Mô hình hồi quy tuyến tính.
- Vec tơ hỗ trợ học máy.
- Phân cụm K-mean



4.2.1 Tổng quan

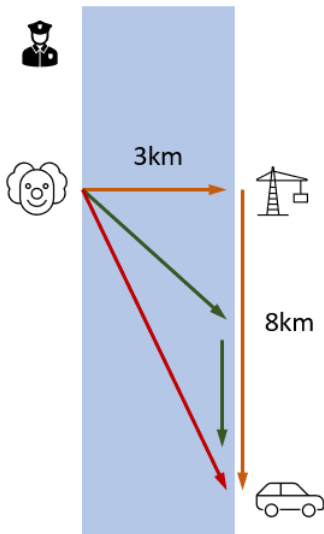
- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
 - chi phí thấp nhất;



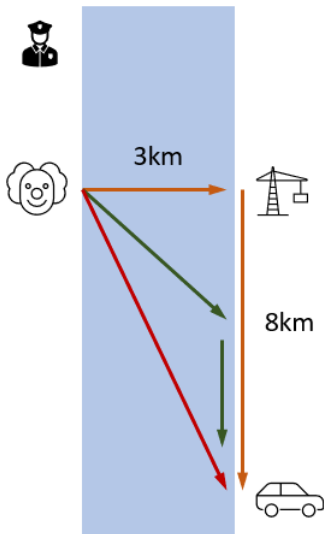


4.2.1 Tổng quan

- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
- chi phí thấp nhất;
 - lợi nhuận cao nhất;

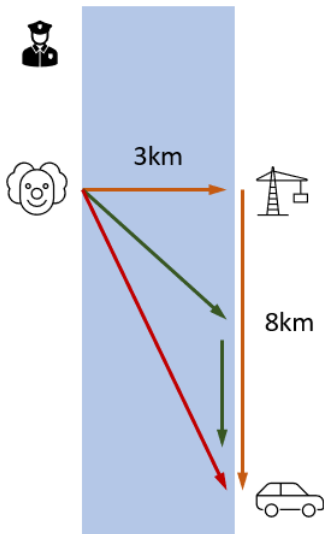


4.2.1 Tổng quan



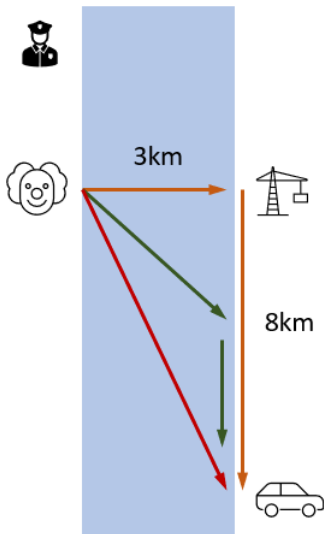
- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
- chi phí thấp nhất;
 - lợi nhuận cao nhất;
 - thời gian ngắn nhất; rủi ro thấp nhất;...

4.2.1 Tổng quan



- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
- chi phí thấp nhất;
 - lợi nhuận cao nhất;
 - thời gian ngắn nhất; rủi ro thấp nhất;...

4.2.1 Tổng quan



- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
 - chi phí thấp nhất;
 - lợi nhuận cao nhất;
 - thời gian ngắn nhất; rủi ro thấp nhất;...

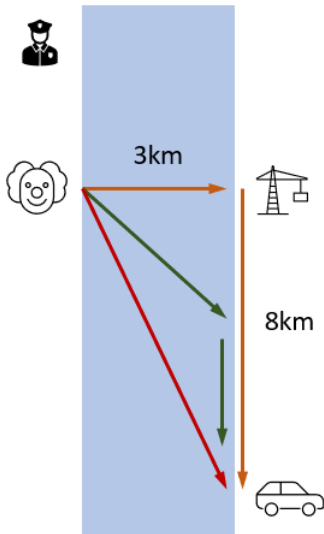
□ Mô hình:

$$\min f(x, y, \dots)$$

$$s.t. (x, y, \dots) \in S \subset \mathbb{R}^n.$$

- x, y, \dots biến số,

4.2.1 Tổng quan



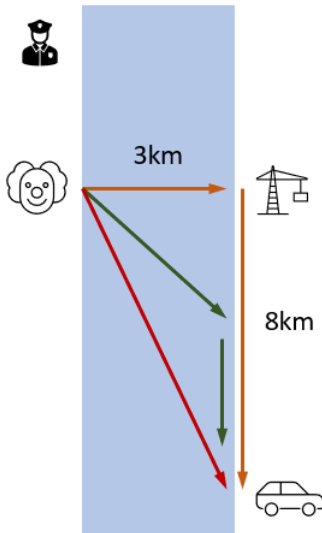
- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
- chi phí thấp nhất;
 - lợi nhuận cao nhất;
 - thời gian ngắn nhất; rủi ro thấp nhất;...

□ Mô hình:

$$\begin{aligned} \min & f(x, y, \dots) \\ \text{s.t. } & (x, y, \dots) \in S \subset \mathbb{R}^n. \end{aligned}$$

- x, y, \dots biến số,
- $f(x, y, \dots)$ hàm mục tiêu,

4.2.1 Tổng quan

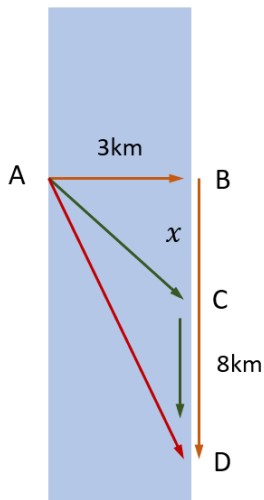


- Tối ưu: Cách hiệu quả nhất để thực hiện "công việc" với
- chi phí thấp nhất;
 - lợi nhuận cao nhất;
 - thời gian ngắn nhất; rủi ro thấp nhất;...

□ Mô hình:

$$\begin{aligned} \min f(x, y, \dots) \\ \text{s.t. } (x, y, \dots) \in S \subset \mathbb{R}^n. \end{aligned}$$

- x, y, \dots biến số,
- $f(x, y, \dots)$ hàm mục tiêu,
- S tập điều kiện (ví dụ: $S = \{(x, y) : g(x, y) = k\}; \dots$)



□ Biến số:

x : Khoảng cách từ C đến D.

□ Mục tiêu tối ưu tổng thời gian để duy chuyển đến ô tô

$$T(x) = \frac{|AC|}{6} + \frac{|CD|}{8} = \frac{\sqrt{x^2 + 9}}{6} + \frac{8 - x}{8}.$$

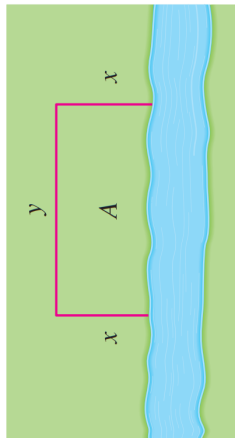
□ Điều kiện:

$$0 \leq x \leq 8.$$

□ Bài toán tối ưu:

$$\begin{aligned} \min \quad & \frac{\sqrt{x^2 + 9}}{6} + \frac{8 - x}{8} \\ \text{s.t.} \quad & 0 \leq x \leq 8. \end{aligned}$$

Bài toán quay rào: một anh nông dân có 2400m dây kẽm. Anh ta muốn quay một bãi chăn bò hình chữ nhật với một mặt là bờ sông sao cho có diện tích lớn nhất.



- ❑ Biến số: x, y : *Chiều dài 2 cạnh.*
- ❑ Mục tiêu: Tối đa diện tích rào được

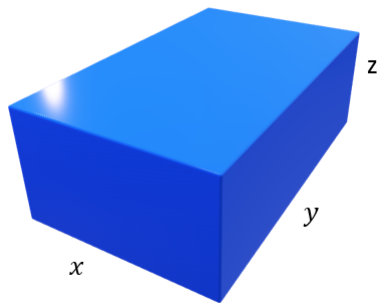
$$A(x, y) = xy.$$

- ❑ Điều kiện: $0 \leq x, y; 2(x + y) = 2400.$
- ❑ Bài toán tối ưu:

$$\min xy$$

$$s.t. 0 \leq x, y; 2(x + y) = 2400.$$

Bài toán container: Một container có sức chứa $10m^3$. Chi phí làm phần đáy là $150/m^2$ và mặt bên là $50/m^2$. Làm thế nào để tối ưu chi phí sản xuất một chiếc container?



- Biến số: x, y, z : Chiều dài 3 cạnh.
- Mục tiêu: Tối thiểu chi phí

$$C(x, y) = 150 * 2 * xy + 50 * (2 * xz + 2 * yz).$$

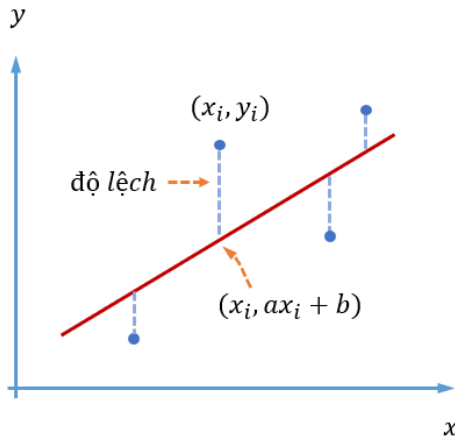
- Điều kiện: $0 \leq x, y, z; xyz = 10$.
- Bài toán tối ưu:

$$\begin{aligned} \min \quad & 300xy + 100xz + 100yz \\ \text{s.t.} \quad & 0 \leq x, y, z; xyz = 10. \end{aligned}$$



4.2.2 Mô hình hồi quy và vec tơ hỗ trợ học máy

Mối quan hệ của chiều cao và cân nặng: Tìm hàm $y = f(x)$ với y : Cân nặng cao (cm) và x : Chiều cao. (kg)



☐ Thực tế: f chỉ có thể ước lượng.

☐ Dự đoán:

$$f(x) = ax + b$$

i.e., quan hệ này là tuyến tính.

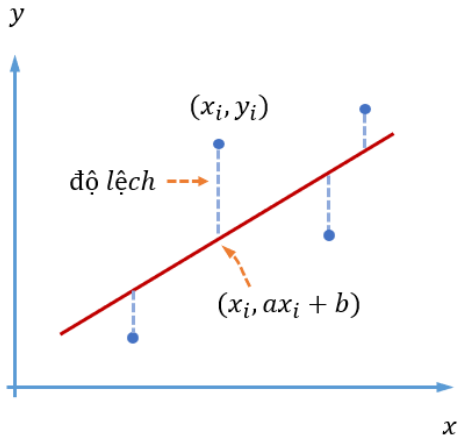
☐ Thu thập dữ liệu n uitters: (x_i, y_i) với $i = \overline{1, n}$.

☐ Mục tiêu: Tìm a, b sao cho đồ thị "fits" data nhất.

☐ Bài toán tối ưu:

$$\min \ell(a, b) = \sum_{i=1}^n \left(y_i - (ax_i + b) \right)^2.$$

☐ Hàm $\ell(a, b)$ còn được gọi là hàm mất mát.



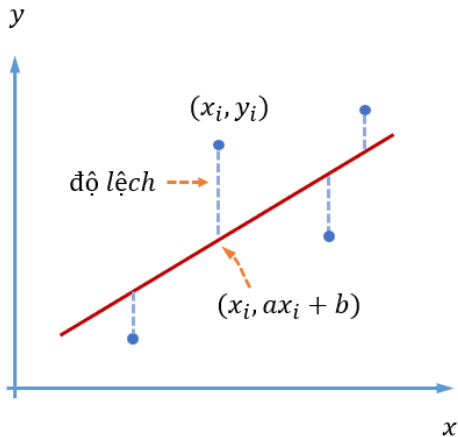
□ Bài toán tối ưu:

$$\min \ell(a, b) = \sum_1^n \left(y_i - (ax_i + b) \right)^2.$$

□ Giải: a, b thỏa điều kiện

$$\frac{\partial \ell}{\partial a} = \sum_1^n 2 \left(y_i - (ax_i + b) \right) (-x_i) = 0$$

$$\frac{\partial \ell}{\partial b} = \sum_1^n 2 \left(y_i - (ax_i + b) \right) (-1) = 0.$$



□ Bài toán tối ưu:

$$\min \ell(a, b) = \sum_1^n \left(y_i - (ax_i + b) \right)^2.$$

□ Giải: a, b thỏa điều kiện

$$a \sum_1^n x_i^2 + b \sum_1^n y_i = \sum_1^n x_i y_i$$

$$a \sum_1^n x_i + nb = \sum_1^n y_i.$$



❑ Tìm sự quan hệ giữa chiều cao và cân nặng; và số sản phẩm bán ra với mức độ đầu tư theo bảng số liệu bên dưới?

Chiều cao	Cân nặng	Sản phẩm bán	Đầu tư (triệu vnd)
1.6	45	1	1
1.65	50	1	2
1.7	55	2	3
1.75	50	2	4
1.75	60	3	5
1.8	60	4	6
1.8	55	6	7



□ Quan hệ giữa cân nặng và chiều cao được cho bởi biểu thức

$$y = ax + b,$$

với a, b là nghiệm của hệ

$$a \sum_{i=1}^7 x_i^2 + b \sum_{i=1}^7 x_i = \sum_{i=1}^7 x_i y_i$$

$$207775 * a + 1205 * b = 64750$$

$$1205 * a + 7 * b = 375.$$

$$a \sum_{i=1}^7 x_i + 7b = \sum_{i=1}^7 y_i.$$

Giải nghiệm ta được

$$y \approx -203 * x + 35014.$$



❑ Dường như kết quả trên không phù hợp với dự đoán? Đây là điểm sai. (Số liệu chưa chuẩn, chưa đủ nhiều để bắt được quy luật tự nhiên).



- ❑ Dường như kết quả trên không phù hợp với dự đoán? Đây là điểm sai. (Số liệu chưa chuẩn, chưa đủ nhiều để bắt được quy luật tự nhiên).
- ❑ Làm thế nào để biết cân nặng phụ thuộc vào chiều cao, chứ không phải ngược lại? (Xác suất thống kê: Hệ số tương quan).



- ❑ Dường như kết quả trên không phù hợp với dự đoán? Đây là điểm sai. (Số liệu chưa chuẩn, chưa đủ nhiều để bắt được quy luật tự nhiên).
- ❑ Làm thế nào để biết cân nặng phụ thuộc vào chiều cao, chứ không phải ngược lại? (Xác suất thống kê: Hệ số tương quan).
- ❑ Nếu y là giá căn hộ chung cư, x_1, x_2 lần lượt biến số chỉ diện tích và vị trí làm thế nào xây dựng mô hình hồi quy cho nhiều biến để dự đoán giá căn hộ:



- ❑ Đường như kết quả trên không phù hợp với dự đoán? Đây là điểm sai. (Số liệu chưa chuẩn, chưa đủ nhiều để bắt được quy luật tự nhiên).
- ❑ Làm thế nào để biết cân nặng phụ thuộc vào chiều cao, chứ không phải ngược lại? (Xác suất thống kê: Hệ số tương quan).
- ❑ Nếu y là giá căn hộ chung cư, x_1, x_2 lần lượt biến số chỉ diện tích và vị trí làm thế nào xây dựng mô hình hồi quy cho nhiều biến để dự đoán giá căn hộ:

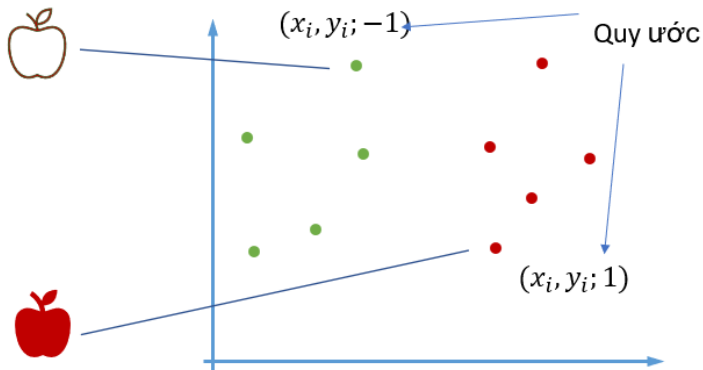
$$y = f(x_1, x_2) = ax_1 + bx_2 + c.$$

- ❑ Mục tiêu: tìm a, b, c sao cho đồ thị "fits" data nhất.
- ❑ Bài toán tối ưu:

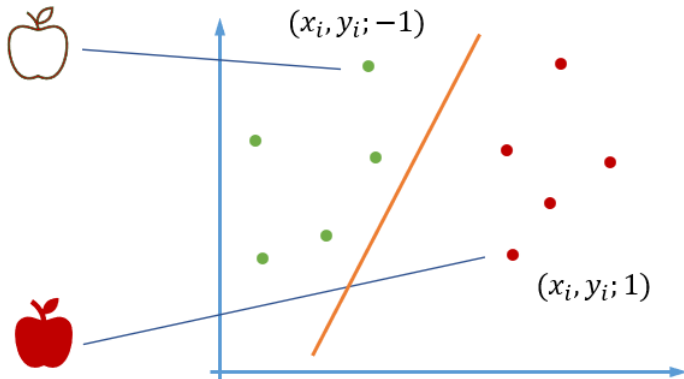
$$\min \ell(a, b, c) \left(:= \sum_{i=1}^n \left(y_i - (ax_{1i} + bx_{2i} + c) \right)^2 \right).$$

- ❑ Hàm $\ell(a, b, c)$ còn được gọi là hàm mất mát.

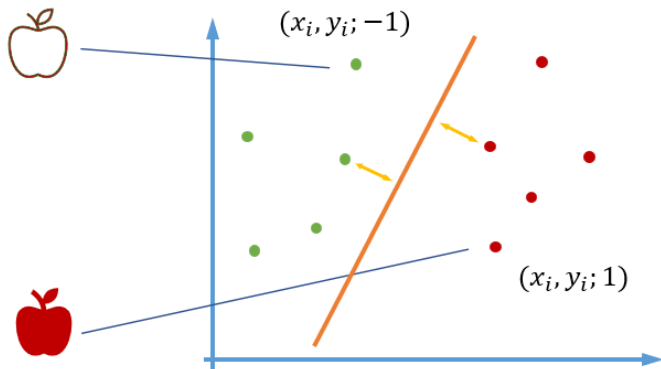
Nhận diện táo: Trong 100 bức ảnh kỹ thuật số có 45 ảnh táo đỏ và 55 táo xanh. Uiters hãy viết chương trình máy tính phân loại táo dựa trên tập dữ liệu trên.



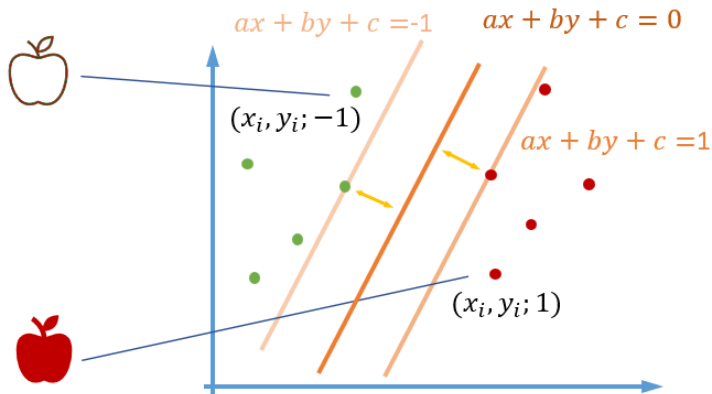
- ❑ Các điểm có chung đặc tính (táo đỏ/xanh) có xu hướng đứng kề nhau.
- ❑ Ý tưởng: Tìm đường thẳng (lựa chọn đơn giản nhất) tách tập điểm thành 2 phần



□ Tối ưu: Tìm đường thẳng sao cho khoảng cách đến 2 tập điểm là bằng nhau và là lớn nhất.



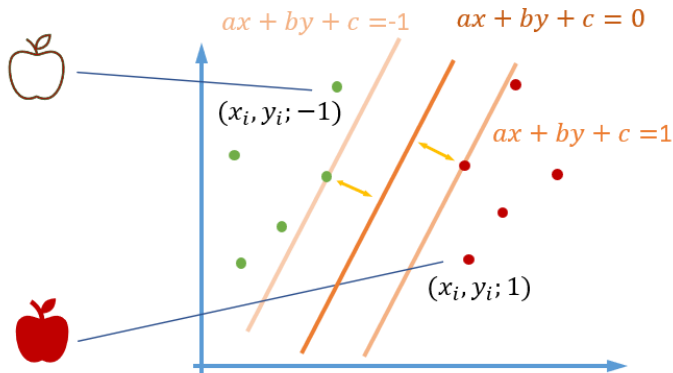
□ Nếu tồn tại $ax + by + c = 0$ cách đến 2 tập điểm bằng nhau. Bằng cách nhân với hệ số thích hợp ta được



□ Bài toán tối ưu:

$$\max d(a, b) = \frac{2}{a^2 + b^2} : \text{ khoảng cách giữa } ax + by + c = \pm 1.$$

□ Điều kiện: $ax_i + by_i + c \geq 1$ nếu $z_i = 1$; và $ax_i + by_i + c \leq -1$ nếu $z_i = -1$;





□ Bài toán tối ưu:

$$\begin{aligned} \max d(a, b) &= \frac{2}{a^2 + b^2} \\ \text{s.t.}, \quad z_i(ax_i + by_i + c) &\geq 1. \end{aligned}$$

□ Bài toán tối ưu:

$$\begin{aligned} \min \quad &\frac{1}{2}(a^2 + b^2) \\ \text{s.t.} \quad &z_i(ax_i + by_i + c) \geq 1. \end{aligned}$$

□ Nếu (x^*, y^*) thông số biểu thị cho một quả táo chưa được đánh dấu, làm sao nhận biết nó là xanh hay đỏ?



□ Bài tập: Táo đỏ $(0, 0), (1, 1), (-1, 1), (1, 2)$; táo xanh $(2, 2), (2, 3), (5, 1)$, viết tường minh bài toán trên?

$$\begin{aligned} \min \quad & \frac{1}{2}(a^2 + b^2) \\ \text{s.t.} \quad & 1 * (a * 0 + b * 0 + c) \geq 1, \\ & 1 * (a * 1 + b * 1 + c) \geq 1, \\ & 1 * (-a * 1 + b * 1 + c) \geq 1, \\ & 1 * (-a * 1 + b * 2 + c) \geq 1, \\ & -1 * (a * 2 + b * 2 + c) \geq 1, \\ & -1 * (a * 2 + b * 3 + c) \geq 1, \\ & -1 * (a * 5 + b * 1 + c) \geq 1, \end{aligned}$$



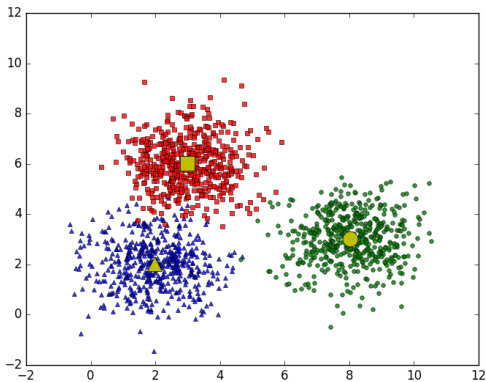
□ Bài tập: Táo đỏ $(0, 0), (1, 1), (-1, 1), (1, 2)$; táo xanh $(2, 2), (2, 3), (5, 1)$, viết tường minh bài toán trên?

$$\begin{aligned} \min \quad & \frac{1}{2}(a^2 + b^2) \\ \text{s.t.} \quad & 1 * (a * 0 + b * 0 + c) \geq 1, \\ & 1 * (a * 1 + b * 1 + c) \geq 1, \\ & 1 * (-a * 1 + b * 1 + c) \geq 1, \\ & 1 * (-a * 1 + b * 2 + c) \geq 1, \\ & -1 * (a * 2 + b * 2 + c) \geq 1, \\ & -1 * (a * 2 + b * 3 + c) \geq 1, \\ & -1 * (a * 5 + b * 1 + c) \geq 1, \end{aligned}$$

□ Tìm nghiệm tối ưu bài toán trên.

Bài toán phân cụm K-means

□ Cho tập dữ liệu gồm N phần tử, $x^i \in \mathbb{R}^d$, $i = \overline{1, N}$. Mục đích của chúng ta phân tách bộ dữ liệu trên thành K tập điểm có "tính chất giống nhau".



Hình: Nguồn: machinelearningcoban.com



□ Với mọi điểm dữ liệu x^i , nếu nó được phân vào lớp cluster k thì ta đặt $y^{ik} = 1$, ngược lại $y^{ik} = 0$.

$$y^{ik} \in \{0, 1\}, \sum_{k=1}^K y^{ik} = 1.$$



□ Với mọi điểm dữ liệu x^i , nếu nó được phân vào lớp cluster k thì ta đặt $y^{ik} = 1$, ngược lại $y^{ik} = 0$.

$$y^{ik} \in \{0, 1\}, \sum_{k=1}^K y^{ik} = 1.$$

□ Gọi m^1, m^2, \dots, m^K là tâm của mỗi cluster. Khoảng cách từ điểm dữ liệu x^i đến m^k

$$d(x^i, m^k) = \left((x_1^i - m_1^k)^2 + \dots + (x_d^i - m_d^k)^2 \right)^{\frac{1}{2}}.$$

□ Với mọi điểm dữ liệu x^i , nếu nó được phân vào lớp cluster k thì ta đặt $y^{ik} = 1$, ngược lại $y^{ik} = 0$.

$$y^{ik} \in \{0, 1\}, \sum_{k=1}^K y^{ik} = 1.$$

□ Gọi m^1, m^2, \dots, m^K là tâm của mỗi cluster. Khoảng cách từ điểm dữ liệu x^i đến m^k

$$d(x^i, m^k) = \left((x_1^i - m_1^k)^2 + \dots + (x_d^i - m_d^k)^2 \right)^{\frac{1}{2}}.$$

□ Giả sử x^i được phân vào cluster m^k thì

$$d(x^i, m^k) = y^{ik} d(x^i, m^k) = \sum_{j=1}^K y^{ij} d(x^i, m^j).$$



□ Để việc phân nhóm có ý nghĩa thì tổng khoảng cách các điểm đến tâm cluster phải là nhỏ nhất:

$$\min \sum_{i=1}^N \sum_{j=1}^K y^{ij} d(x^i, m^j)$$
$$s.t. \quad y^{ik} \in \{0, 1\}, \quad \sum_{k=1}^K y^{ik} = 1.$$