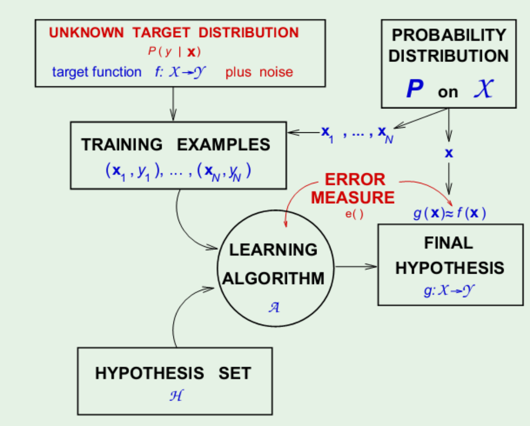
Компьютерное зрение ’2014
Машины опорных векторов

Who? Александр Вахитов

When? May 2, 2014

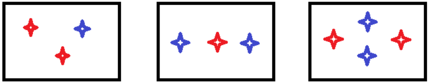
Прошлые лекции (1)



(c) Y. Abu-Mostafa, Learning From Data (Caltech online)

Прошлые лекции (2)

Функция роста $m_H(N)$: сколько вариантов классификации N точек данным множеством гипотез
Размерность VC (Вапник-Червоненкис): максимальное число точек, для которых данное множество гипотез способно породить все возможные варианты бинарной классификации точек (дихотомии)
Пример: разделяющие прямые на плоскости, бинарная классификация



$$m_H(3) = 8 = 2^3 \quad m_H(4) = 14 < 2^4$$

Прошлые лекции (3)

PCA - метод главных компонент
Пример: распознавание лиц - размерность 256 на 256, в то время как "смысловая" размерность ниже (тип носа, глаз, губ, ушей...)
Метод: фиксируем целевую размерность s, выбираем s наиболее осмысленных координат (признак - вариативность по координате), выполняем классификацию в выбранном подпространстве размерности s

О чем мы поговорим

- баланс сложности и точности гипотезы (bias-variance tradeoff)
- штраф за излишнюю сложность: регуляризация
- машины опорных векторов

Баланс сложности и точности

g_D - гипотеза, выбранная по набору данных
$D = (x_1, y_1) \dots (x_N, y_N)$
$\bar{g} = E_D g_D$ - средняя гипотеза (мат. ожидание по набору данных)
$f(x)$ искомая неизвестная функция
x некий входной вектор

$$E_x E_D (f(x) - g_D(x))^2 = E_x E_D (f(x) - \bar{g}(x) + \bar{g}(x) - g_D(x))^2$$

$$E_x E_D (f(x) - g_D(x))^2 = E_x (f(x) - \bar{g}(x))^2 + E_x E_D (\bar{g}(x) - g_D$$

точность: $E_x (f(x) - \bar{g}(x))^2$
сложность(вариативность) гипотез:
$E_x E_D (\bar{g}(x) - g_D(x))^2$

Баланс сложности и точности: пример

точность: $E_x(f(x) - \bar{g}(x))^2$
сложность(вариативность) гипотез:
$E_x E_D(\bar{g}(x) - g_D(x))^2$

$$f(x) = cx + d + v,\ x, c, d \in \mathbb{R}$$

v - случайный шум 2 точки, гипотеза - константа: (точность: гипотеза проще реальности, есть избыточность для имеющегося набора данных)
2 точки, гипотеза - прямая: ошибочна из-за шума (точность адекватна реальности, но сложность велика для имеющегося набора данных)
3 точки, Гипотеза - прямая: ближе к реальности (точность адекватна реальности, есть избыточность)

Регуляризация

$$Ax = b, x \in \mathbb{R}^q, b \in \mathbb{R}^N, A \in \mathbb{R}^q \times \mathbb{R}^N,\ q > N$$

x определен неоднозначно
Рассмотрим функцию $R(x) = \|Ax - b\|^2 + \lambda\|x\|^2$, $\lambda > 0$

$$R(x) = \|Ax - b\|^2 + \lambda\|x\|^2 \to \min$$

Минимум уникален (для $x \in b + KerA$ $R(x) = \lambda\|x\|^2$)
Впервые подход предложен А.Н. Тихоновым

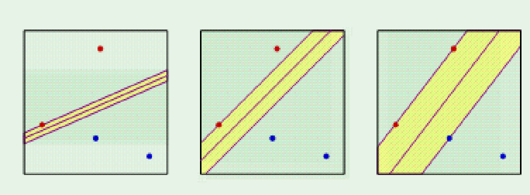
Регуляризация в машинном обучении

Есть множество гипотез H. Введем регуляризационную функцию - штраф за "излишнюю" сложность
Пример: точки на плоскости, гипотеза - прямая $h(x) = h^T x + b$, регуляризационный член $\lambda(\|h\|^2 + d^2)$ штраф за отклонение от оси OX
Вопрос: как выбирать λ?

Машины опорных векторов (Support Vector Machines - SVM)

Задача бинарной классификации на плоскости разделяющими прямыми. Три точки.

$$y = w^T x + b$$

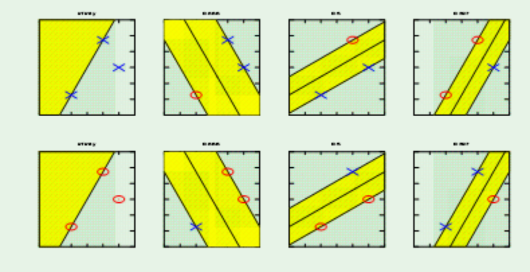


1. Почему чем больше поле (желтое), тем лучше?
2. Какое значение w, b максимизирует жетое поле?

Смысл максимизации размера поля

1. Прямая, позволяющая наибольшее возможное отклонения без ошибок классификации на тренировочной выборке
2. Требование величины поля большей, чем ϵ, может снизить значение функции роста $m_H(N)$ и облегчить обучение

Требование необходимой величины поля снижает функцию роста



Вывод формулы

Пусть x_n ближайшая к плоскости точка.
Расстояние $dist(x_n)$ от точки x_nдо гиперплоскости w, b равно $\frac{1}{\|w\|} w^T x_n$.
Выберем w так, чтобы $|w^T x_n + b| = 1$.
Тогда

$$dist(x_n) = \frac{1}{\|w\|}$$

Задача оптимизации 1

$$\frac{1}{\|w\|} \to \max \;\; s.t. \min_{n=1,2,\dots,N} |w^T x_n + b| = 1$$

Заметим, что $|w^T x_n + b| = y_n(w^T x_n + b)$
Переформулируем:

$$\frac{1}{2} w^T w \to \min \;\; s.t. \; y_n(w^T x_n + b) \geq 1 \;\; n = 1, 2, \dots, N$$

Формулировка в виде лагранжиана

Экстремальная задача с ограничениями: условия Куна-Такера

$$L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{n=1}^{N} \alpha_n (y_n (w^T x_n + b) - 1)$$

Минимизировать $L(w, b, \alpha)$ по w, b и максимизировать по $\alpha \geq 0$

$$\nabla_w L = w - \sum_{n=1}^{N} \alpha_n y_n x_n = 0$$

$$= -\sum_{n=1}^{N} \alpha_n y_n = 0$$

Подстановка

Подставим $w = \sum \alpha_n y_n x_n$, $\sum_{n=1}^{N} \alpha_n y_n = 0$ в

$$L(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{n=1}^{N} \alpha_n (y_n (w^T x_n + b) - 1)$$

получим

$$L(\alpha) = \sum_{n=1}^{N} \alpha_n - \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} y_n y_m \alpha_n \alpha_m x_n^T x_m$$

максимизировать при условии $\alpha_n \geq 0$ для $n = 1, 2, \ldots, N$ и $\sum_{n=1}^{N} \alpha_n y_n = 0$

Квадратичное программирование

$$\min \frac{1}{2}\alpha^T Q\alpha - \mathbf{1}^T\alpha_n,\ Q_{ij} = y_i y_j x_i^T x_j$$

при $y^T\alpha = 0$, $\alpha \geq 0$

Решение: α^*

$$w^* = \sum_{n=1}^{N} \alpha_n^* y_n x_n$$

При $\alpha_n > 0$ x_n называется опорным вектором

Любой опорный вектор позволяет вычислить

$b^* = y_n - w^T x_n$

Ожидаемая ошибка для SVM

Обозначим число опорных векторов как s.

$$P(h_{SVM}(x) \neq f(x)) \leq \frac{Es}{N-1}$$

Задание по SVM

Классификация рукописных цифр

1. скачиваем базу NIST
2. считаем HOG для каждой цифры
3. запускаем svm light

Задание по проективной геометрии (1)

Одна камера, K - единичная ($P = [Rt]$)

$$x \equiv PX = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} X = \begin{pmatrix} p_1 X \\ p_2 X \\ p_3 X \end{pmatrix}$$

1. какой смысл у p_i?
2. как p_i связаны между собой?
3. что является решением $p_i X = 0$, $i = 1 \dots 3$?

Задание по проективной геометрии (2)

Известно, что

$$K = \begin{pmatrix} 500 & 0 & 800 \\ 0 & 500 & 800 \\ 0 & 0 & 1 \end{pmatrix}$$

Кадр, сделанный этой камерой, имеет размер 1600 на 1600 пикселей.
Пусть известно, что точка схода рельсов имеет координаты (1507, 1300), а точка схода шпал - координаты (800, 300).
Определить углол обзора камеры и углы между

1. осью камеры и плоскостью
2. осью камеры и рельсами

Задание по проективной геометрии (3)

Известно, что

$$K = \begin{pmatrix} 500 & 0 & 800 \\ 0 & 500 & 800 \\ 0 & 0 & 1 \end{pmatrix}$$

Кадр, сделанный этой камерой, имеет размер 1600 на 1600 пикселей.

Пусть человек стоит на плоской поверхности, камера направлена параллельно поверхности, на высоте 2 м. Сделан кадр, затем человек отошел вправо и повернулся налево на 45 градусов. Сделал второй кадр.

Точка находится на расстоянии 10 м по направлению камеры от первой точки съемки и на высоте 4 м над поверхностью. Определить

1. ее проекцию на первую и вторую камеры
2. уравнение эпиполярной линии, соответствующей проекции точки на первую камеру, на второй камере