CS246 (Winter 2015) Mining Massive Data Sets

Introduction to proof techniques

Sammy El Ghazzal, Jean-Yves Stephan

Disclaimer These notes may contain typos, mistakes or confusing points. Please contact jstephan@stanford.edu so that we can improve them for next year.

1 Proof by induction

The proof by induction is usually illustrated by falling dominoes. Make the first domino fall (initialization or base case) and if your dominos are placed in such a way that each domino makes the next one fall (inheritance or induction), then all dominos will be down.

1.1 Principle

Say you want to prove that a property holds for all integers. To do a proof by induction, you will:

1. Base case: Prove that the property holds for $n = 0$ (possibly another value, it really depends on the cases).

2. Inheritance: Assume that the result holds[1] for $n \geq 0$ and prove that the property would hold for $n+1$.

1.2 Example

1. Let us start with a simple example. Prove that:

$$\forall n \geq 1,\ \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

Solution We proceed in two steps:

- Base case ($n = 1$): the left hand side is 1, as is the right hand side so the property holds for $n = 1$.
- Now, we assume that the property holds for a given $n \geq 1$, and we want to prove that it holds for $n + 1$.

[1] There is a stronger variant of the proof by induction where you assume that the property holds for any integer $k \leq n$.

We compute:

$$\begin{aligned}
\sum_{i=1}^{n+1} i^2 &= \sum_{i=1}^{n} i^2 + (n+1)^2 \\
&= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \qquad \text{(Induction hypothesis)} \\
&= \frac{n+1}{6}\left(n(2n+1) + 6(n+1)\right) \\
&= \frac{n+1}{6}\underbrace{\left(2n^2+7n+6\right)}_{(n+2)(2n+3)} \\
&= \frac{(n+1)(n+2)(2n+3)}{6} \\
&= \frac{(n+1)((n+1)+1)(2(n+1)+1)}{6},
\end{aligned}$$

which proves that the property holds for $n+1$ and concludes the proof by induction.

2. A more complicated example[2]: let $n \geq 1$, and two families of vector $(e_1, \dots, e_n)$, and $(f_1, \dots, f_{n+1})$ such that:

$$\forall 1 \leq j \leq n+1,\ f_j = \sum_{i=1}^{n} \lambda_i^{(j)} e_i.$$

Prove that $(f_1, \dots, f_{n+1})$ is a dependent[3] family.

Solution Let us denote by $H(n)$ the property that we want to prove by induction:
$H(n) =$ For any families $f = (f_1, \dots, f_{n+1})$ and $e = (e_1, \dots, e_n)$ such that:

$$\forall 1 \leq j \leq n+1,\ f_j = \sum_{i=1}^{n} \lambda_i^{(j)} e_i,$$

then f is dependent.

We proceed in two steps:

- Base case ($n = 1$): we have:

$$\exists \lambda_1, \lambda_2,\ f_1 = \lambda_1 e_1,\ f_2 = \lambda_2 e_1.$$

If both λ_1 and λ_2 are zero, then the family is clearly dependent.
Otherwise, we compute:

$$\lambda_1 e_2 - \lambda_2 e_1 = 0,$$

which proves that the family (e_1, e_2) is dependent.

- Inheritance: Assume that $H(n)$ holds and let us try to prove that $H(n+1)$ holds.
We distinguish two cases:

[2] Note that the following result basically says that any set of $n+1$ vectors in a vector space of dimension n is always dependent.

[3] Recall that $(v_1, \dots, v_n)$ is a (linearly) dependent family if and only if:

$$\exists (\lambda_1, \dots, \lambda_n) \neq 0,\ \sum_{i=1}^{n} \lambda_i v_i = 0.$$

(a) Forall $1 \leq j \leq n+2$, $\lambda_{n+1}^{(j)} = 0$. Then, we can rewrite all the f_j's in terms of $(e_1, \ldots, e_n)$ and therefore using $H(n)$ we conclude that $(f_1, \ldots, f_{n+1})$ is dependent and so will be $(f_1, \ldots, f_{n+2})$.

(b) One of the $\lambda_{n+1}^{(j)}$ is non-zero. We can assume (by changing the numbering if necessary) that $\lambda_{n+1}^{(n+2)} \neq 0$. We then define:

$$\forall 1 \leq j \leq n+1,\ \tilde{f}_j = f_j - \frac{\lambda_{n+1}^{(j)}}{\lambda_{n+1}^{(n+2)}} f_{n+2}.$$

Then, the coefficient on e_{n+1} is zero for all the $\tilde{f}_j$'s, and therefore, we can apply $H(n)$ to $(\tilde{f}_1, \ldots, \tilde{f}_{n+1})$ and $(e_1, \ldots, e_n)$, which gives us that $(\tilde{f}_1, \ldots, \tilde{f}_{n+1})$ is dependent, that is:

$$\exists (\beta_1, \ldots, \beta_{n+1}) \neq 0,\ \beta_1 \tilde{f}_1 + \cdots + \beta_{n+1} \tilde{f}_{n+1} = 0,$$

which can be rewritten:

$$\exists (\beta_1, \ldots, \beta_{n+1}) \neq 0,\ \beta_1 f_1 + \cdots + \beta_{n+1} f_{n+1} - \sum_{j=1}^{n+1} \frac{\beta_j \lambda_{n+1}^{(j)}}{\lambda_{n+1}^{(n+2)}} f_{n+2} = 0,$$

which by definition means that $(f_1, \ldots, f_{n+2})$ is dependent and concludes the inheritance part of the proof.

1.3 Common mistakes

- Forgetting the base case. Usually, the base case is easy compared to the induction step, but you still have to clearly mention that you looked at it.

2 Proof by contrapositive

2.1 Principle

Say you want to prove something of the form:

$$A \Rightarrow B.$$

The usual way to prove this is to assume that A is true and then try to get to B. Sometimes it is easier to prove that:

$$\neg B \Rightarrow \neg A.$$

2.2 Example

1. Prove that for any $a, b \in \mathbb{R}$:

$$a + b \text{ is irrational} \Rightarrow a \text{ or } b \text{ is irrational}.$$

Solution $\neg B$ in our case will be:

$$a \text{ and } b \text{ are rational.}$$

Now we want to prove $\neg A$, that is:

$$a+b \text{ is rational.}$$

This is now straightforward: indeed, we can write:

$$\exists p, q, u, v \in \mathbb{N},\ a = \frac{p}{q} \text{ and } b = \frac{u}{v}.$$

We then compute:

$$a + b = \frac{pv + qu}{qv},$$

which proves that $a+b$ is rational and concludes the proof by contrapositive.

2. Prove that for $n \in \mathbb{N}$:

$$8 \text{ does not divide } n^2 - 1 \Rightarrow n \text{ is even.}$$

Hint: Notice that any odd number n can be written as $n = 4k + r$ with $k \in \mathbb{N}$ and $r \in \{1, 3\}$.

Solution The contrapositive is:

$$n \text{ is odd} \Rightarrow 8 \text{ divides } n^2 - 1.$$

Now, let us do the proof: let us take n an odd integer. By using the hint, we write n as $4k + r$ with $r \in \{1, 3\}$.

We compute:

$$n^2 - 1 = 16k^2 + 8rk + r^2 - 1.$$

Now by noticing that 8 divides $16k^2$, $8rk$ and $r^2 - 1$ (because this quantity is either 0 or 8), we conclude that 8 divides $n^2 - 1$ which is what we wanted.

2.3 Common mistakes

- Computing $\neg A$ and $\neg B$ may be tricky in some cases. You should spend some time practicing if you are not confident on this point.

  For instance, try to find the contrapositive of the following (definition of $\lim_{x \to a} f(x) = b$):

$$\forall \epsilon > 0,\ \exists \alpha > 0, \forall x\ ||x - a|| \leq \alpha \Rightarrow ||f(x) - b|| \leq \epsilon$$

Solution The contrapositive is:

$$\exists \epsilon > 0,\ \forall \alpha > 0, \exists x \text{ s.t. }\ ||x - a|| \leq \alpha \text{ and } ||f(x) - b|| > \epsilon$$

3 Proof by contradiction

3.1 Principle

A proof by contradiction works as follows: assume that the property you are trying to prove is wrong and find a contradiction.

3.2 Example

1. Prove that $\sqrt{2}$ is irrational.

   Solution Assume by contradiction that $\sqrt{2}$ is rational, that is[4]:

   $$\exists p, q \text{ such that } \gcd(p,q) = 1 \text{ and } \sqrt{2} = \frac{p}{q}.$$

   Now, our goal is to find a contradiction.

   We compute:

   $$p^2 = 2q^2 \Rightarrow p^2 \text{ is even} \Rightarrow p \text{ is even.}$$

   Therefore, we can write $p = 2k$ and:

   $$2 = \frac{(2k)^2}{q^2} \Leftrightarrow q^2 = 2k^2 \Rightarrow q \text{ is even.}$$

   Because both p and q are even, we have $\gcd(p,q) \geq 2$, which contredicts our assumption that $\gcd(p,q) = 1$.

   Therefore $\sqrt{2}$ is irrational.

2. Show that any function k-Lipschitz with $k < 1$ has at most one fixed point[5].

   Solution Let f be a k-Lipschitz function.

   Assume by contradiction that f has 2 distinct fixed points that we call u and v with $v \neq u$. Then:

   $$\underbrace{||f(u) - f(v)||}_{=||u-v||} \leq k||u - v|| \Leftrightarrow (1-k)||u - v|| \leq 0.$$

   But $1 - k > 0$ and $||u - v|| > 0$ because u and v are distinct. This is a contradiction and therefore our initial assumption is wrong, that is f has at most one fixed point.

---

[4]The fact that we can assume $\gcd(p,q) = 1$ comes from the fact that you can reduce a fraction to its minimal form.

[5]Recall that f is k-Lipschitz on $\mathbb{R}$ if and only if:

$$\forall x, y \in \mathbb{R},\ ||f(x) - f(y)|| \leq k||x - y||.$$