



# *Common Inequalities in Computer Science*

Sambuddha Roy, LinkedIn

July 6, 2016



# *Contents*

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..



# Contents

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..
  - ▶ Approximation Algorithms



# Contents

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..
  - ▶ Approximation Algorithms
  - ▶ Optimization



# Contents

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..
  - ▶ Approximation Algorithms
  - ▶ Optimization
  - ▶ Machine Learning



# Contents

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..
  - ▶ Approximation Algorithms
  - ▶ Optimization
  - ▶ Machine Learning
  - ▶ Mathematical Analysis



# Contents

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..
  - ▶ Approximation Algorithms
  - ▶ Optimization
  - ▶ Machine Learning
  - ▶ Mathematical Analysis
  - ▶ ...



# Contents

---

- ▶ Why inequalities?
- ▶ Used in: almost everywhere..
  - ▶ Approximation Algorithms
  - ▶ Optimization
  - ▶ Machine Learning
  - ▶ Mathematical Analysis
  - ▶ ...
  - ▶ (Can almost be called the backbone of mathematics..)





## *See it in real life*

---

- ▶ So you have an algorithm - how do you prove it to be optimal (or close to optimal)?



## *See it in real life*

---

- ▶ So you have an algorithm - how do you prove it to be optimal (or close to optimal)?
- ▶ You have to show that there is a *lower bound* for the resources that the algorithm takes.



## See it in real life

---

- ▶ So you have an algorithm - how do you prove it to be optimal (or close to optimal)?
- ▶ You have to show that there is a *lower bound* for the resources that the algorithm takes.
- ▶ Essentially - prove inequalities!



## *A simple inequality*

---

- ▶ Setting: We are given  $n$  positive real numbers  $x_1, x_2, \dots, x_n$ , such that:  $x_1 + x_2 + \dots + x_n \geq n$ .
- ▶ Prove:  $\sum_{i: x_i > 1/2} x_i \geq n/2$ .



## *A simple inequality*

---

- ▶ Break up the sum  $x_1 + x_2 + \cdots + x_n$  into two parts, collecting all terms  $i$  such that  $x_i > 1/2$  and terms  $i$  such that  $x_i \leq 1/2$ .

$$\sum_i x_i = \sum_{i: x_i \leq 1/2} x_i + \sum_{i: x_i > 1/2} x_i$$

- ▶ Let  $k$  denote the number of terms such that  $x_i \leq 1/2$ , so that:

$$n \leq \sum_i x_i \leq k/2 + \sum_{i: x_i > 1/2} x_i$$



## A simple inequality

---

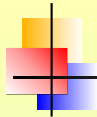
- ▶ Break up the sum  $x_1 + x_2 + \dots + x_n$  into two parts, collecting all terms  $i$  such that  $x_i > 1/2$  and terms  $i$  such that  $x_i \leq 1/2$ .

$$\sum_i x_i = \sum_{i: x_i \leq 1/2} x_i + \sum_{i: x_i > 1/2} x_i$$

- ▶ Let  $k$  denote the number of terms such that  $x_i \leq 1/2$ , so that:

$$n \leq \sum_i x_i \leq k/2 + \sum_{i: x_i > 1/2} x_i$$

- ▶ So that:  $\sum_{i: x_i > 1/2} x_i \geq (n - k/2) \geq n/2$ .



## *A simple inequality: Summary*

---

- ▶ Given the  $n$  numbers, the **mean** is  $\geq 1$ .
- ▶ Maybe very few numbers have  $x_i \approx 1 \dots$  but



## A simple inequality: Summary

---

- ▶ Given the  $n$  numbers, the **mean** is  $\geq 1$ .
- ▶ Maybe very few numbers have  $x_i \approx 1 \dots$  but
- ▶ Most of the **mass** is *around* 1:

$$\sum_{i: x_i > 1/2} x_i \geq n/2$$





## *A simple inequality: Summary*

---

- ▶ Given the  $n$  numbers, the **mean** is  $\geq 1$ .
- ▶ Maybe very few numbers have  $x_i \approx 1 \dots$  but
- ▶ Most of the **mass** is *around* 1:

$$\sum_{i: x_i > 1/2} x_i \geq n/2$$

- ▶ A concentration of measure result.



## Inequality 2

---

- ▶ Setting:
  - ▶ We are given a *digraph*  $D = (V, A)$ .
  - ▶ Given a vertex  $v$ , let  $i(v)$  denote the indegree of the vertex, and let  $d(v)$  denote the total degree (in + out).
  - ▶ Let  $V_1 = \{v : i(v) \geq d(v)/3\}$
- ▶ Required to prove that:

$$\sum_{v \in V_1} d(v) \geq |A|/3.$$



## Inequality 2

---

- ▶ Another application of concentration of measure.
- ▶ Collect facts:
  - ▶  $\sum_{v \in V} d(v) = 2|A|$
  - ▶  $\sum_{v \in V} i(v) = |A|$
- ▶ Rewrite the blue inequality above as
$$\sum_{v \in V} d(v) \cdot \frac{i(v)}{d(v)} = |A|.$$
- ▶ From here, we want to get:  $\sum_{v \in V_1} d(v) \geq |A|/3.$



## Inequality 2: Use blue fact

### Facts

- ▶  $\sum_{v \in V} d(v) = 2|A|$
- ▶  $\sum_{v \in V} i(v) = |A|$
- ▶ Have:  $\sum_{v \in V} d(v) \cdot \frac{i(v)}{d(v)} = |A|$
- ▶ Want to have:  $\sum_{v \in V_1} d(v) \geq |A|/3$ .
- ▶ Separate out the vertices as  $v \in V_1$  and  $v \in V \setminus V_1$ .
- ▶ For  $v \in V \setminus V_1$ ,  $i(v)/d(v) < 1/3$ . For  $v \in V_1$ ,  $i(v)/d(v) \leq 1$ .
- ▶  $|A| = \sum_{v \in V} d(v) \cdot \frac{i(v)}{d(v)} \leq \sum_{v \in V_1} d(v) + \sum_{v \in V \setminus V_1} d(v)/3$



## Inequality 2: Use *red* fact

### Facts

- ▶  $\sum_{v \in V} d(v) = 2|A|$
- ▶  $\sum_{v \in V} i(v) = |A|$
- ▶  $|A| = \sum_{v \in V} d(v) \cdot \frac{i(v)}{d(v)} \leq \sum_{v \in V_1} d(v) + \sum_{v \in V \setminus V_1} d(v)/3$
- ▶ i.e.  $|A| \leq \sum_{v \in V_1} d(v) + \frac{2|A|}{3}$ , which gives
- ▶  $\sum_{v \in V_1} d(v) \geq \frac{|A|}{3}$ .



## *Inequality 2: Quiz Problem 1 (Extra Credit)*

---

- ▶ We showed:

$$\sum_{v \in V_1} d(v) \geq |A|/3.$$



## *Inequality 2: Quiz Problem 1 (Extra Credit)*

---

- ▶ We showed:

$$\sum_{v \in V_1} d(v) \geq |A|/3.$$

- ▶ Instead show that:

$$\sum_{v \in V_1} d(v) \geq |A|/2.$$



## *Inequality 2: Quiz Problem 1 (Bonus Credit)*

---

- Show that:

$$\sum_{v \in V_1} d(v) \geq |A|$$





- ▶ Why is this a “concentration of measure” result?
- ▶ Note that  $\sum_{v \in V} d(v) = 2 \sum_{v \in V} i(v)$ , so *on average*  $i(v)$  is roughly **half** of  $d(v)$ .
- ▶ As before, we move slightly away from the mean and ask: how many vertices  $v$  will satisfy  $i(v) \geq d(v)/3$ ?
- ▶ While we don't get a *count*, we get a result for the *mass*.



## *Concentration of measure*

---

- ▶ Chebyshev Inequalities
- ▶ Markov Inequalities
- ▶ Chernoff-Hoeffding bounds.



## Quiz Problem 2

---

- ▶ Recall Chebyshev's inequality:
  - ▶ Setting: let  $X$  be a random variable with  $E(X) = \mu$  and variance  $\sigma^2 = \text{var}(X)$ .
  - ▶ Then:  $P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$



## Quiz Problem 2

---

- ▶ Recall Chebyshev's inequality:
  - ▶ Setting: let  $X$  be a random variable with  $E(X) = \mu$  and variance  $\sigma^2 = \text{var}(X)$ .
  - ▶ Then:  $P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$
  - ▶ Question: A one-sided version of Chebyshev



## Quiz Problem 2

---

- ▶ Recall Chebyshev's inequality:
  - ▶ Setting: let  $X$  be a random variable with  $E(X) = \mu$  and variance  $\sigma^2 = \text{var}(X)$ .
  - ▶ Then:  $P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$
  - ▶ Question: A one-sided version of Chebyshev
  - ▶ Prove that:

$$P(X - \mu \geq t) \leq \frac{\sigma^2}{\sigma^2 + t^2}.$$



## *Shifting Gears: Lagrangian Relaxations*

---

- ▶ Suppose we are given an optimization problem.
- ▶ What is an optimization problem?
  - ▶ Has an **objective**
  - ▶ Has **constraints**



## *Shifting Gears: Lagrangian Relaxations*

---

- ▶ Suppose we are given an optimization problem.
- ▶ What is an optimization problem?
  - ▶ Has an **objective**
  - ▶ Has **constraints**
- ▶ Some constraints are easy to satisfy



## *Shifting Gears: Lagrangian Relaxations*

---

- ▶ Suppose we are given an optimization problem.
- ▶ What is an optimization problem?
  - ▶ Has an **objective**
  - ▶ Has **constraints**
- ▶ Some constraints are easy to satisfy (called “soft” constraints)





## *Shifting Gears: Lagrangian Relaxations*

---

- ▶ Suppose we are given an optimization problem.
- ▶ What is an optimization problem?
  - ▶ Has an **objective**
  - ▶ Has **constraints**
- ▶ Some constraints are easy to satisfy (called “soft” constraints)
- ▶ while some others are harder - hard constraints.



## Lagrangian Relaxations

---

- ▶ The big idea behind Lagrangian Relaxations is to “relax” the hard constraints by *pulling them* into the objective.



## *An example*

---

An optimization problem:

$$\min \sum_{i \in V} c_i x_i$$

$$\forall (i, j) \in E \quad x_i + x_j \geq 1$$

$$\forall i \in V \quad x_i \geq 0$$

Looks familiar?



## Example modified

Modified optimization problem:

$$\min \sum_{i \in V} c_i x_i$$

$$\forall (i, j) \in E \quad x_i + x_j \geq 1$$

$$\sum_{i \in V} x_i \geq k$$

$$\forall i \in V \quad x_i \geq 0$$

Here, the  $x_i + x_j \geq 1$  are the “soft” constraints, and  $\sum_{i \in V} x_i \geq k$  is the “hard” constraint.



Optimization problem  $O_1$ :

$$\min f(x)$$

s.t.

$$s(x) \leq 0$$

$$h(x) \leq 0$$

where,  $s(x)$  stands for the soft constraints, and  $h(x)$  for the hard constraints.



## Lagrangian Relaxation

---

Relaxation  $O_2 = \text{LR}(\lambda)$  of  $O_1$ :

$$\min f(x) + \lambda \cdot h(x)$$

s.t.

$$s(x) \leq 0$$

$$\lambda \geq 0$$

Here,  $\lambda \in [0, \infty)$  is called a Lagrange Multiplier.

Major Impact:  $\text{LR}(\lambda) \leq O_1 !!!$



## Hölder's Inequality

---

Setting:

- ▶ Vectors  $\mathbf{x} = (x_1, \dots, x_n)$ ,  $\mathbf{y} = (y_1, \dots, y_n)$ , where  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_{\geq 0}^n$ .
- ▶ Reals  $p, q \in \mathbb{R}$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ .

Prove that:

$$(x_1^p + x_2^p + \dots + x_n^p)^{1/p} \cdot (y_1^q + y_2^q + \dots + y_n^q)^{1/q} \geq \mathbf{x} \cdot \mathbf{y}$$



## Hölder's Inequality

---

Have we seen (any special case of) this before? With  $p = q = 2$ :

$$(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} \cdot (y_1^2 + y_2^2 + \cdots + y_n^2)^{1/2} \geq \mathbf{x} \cdot \mathbf{y}$$

which is the Cauchy Schwarz inequality!





## Another Hölder's Inequality

---

The following is also often called Hölder's Inequality (same conditions on  $p, q$ , non-negative real numbers  $x, y$ ):

$$\frac{x^p}{p} + \frac{y^q}{q} \geq xy$$

Proof is easy by AM – GM. (Quiz Problem 3). Call this the “little” Hölder's Inequality. Condition for equality:  $x^p = y^q$ . Now, from here to the first Hölder's Inequality? We will only discuss  $n = 2$ .



$$n = 2$$

We want:

$$(x_1^p + x_2^p)^{1/p} \cdot (y_1^q + y_2^q)^{1/q} \geq \mathbf{x} \cdot \mathbf{y}$$

First attempt: We can write down 2 “little” Hölder Inequalities (one for  $x_1, y_1$  another for  $x_2, y_2$ ) but that gives us:

$$\frac{x_1^p + x_2^p}{p} + \frac{y_1^q + y_2^q}{q} \geq \mathbf{x} \cdot \mathbf{y}$$

This is **weaker** than what we want; why?



## Weaker?

---

So far, we have:

$$\frac{x_1^p + x_2^p}{p} + \frac{y_1^q + y_2^q}{q} \geq \mathbf{x \cdot y}$$

Let  $z_1^p = x_1^p + x_2^p$  and  $z_2^q = y_1^q + y_2^q$ , then one application of the “little” Hölder Inequality gives:

$$\frac{x_1^p + x_2^p}{p} + \frac{y_1^q + y_2^q}{q} \geq z_1 \cdot z_2 = (x_1^p + x_2^p)^{1/p} \cdot (y_1^q + y_2^q)^{1/q}$$



## *Weaker to Stronger*

---

Terry Tao's suggestion: in order to milk most out of an inequality, apply it to a scenario where equality can even conceivably hold.

Also called the “principle of maximum effectiveness”.

We realize that for the application of the “little” Hölder Inequality above, the equality condition of  $x_1^p + x_2^p = y_1^q + y_2^q$  may be widely flouted.



## Enter Lagrangian Multipliers

---

Taking the cue from there, let's apply a Lagrangian Multiplier  $\lambda$ :

$$\frac{x_1^p + x_2^p}{p} + \lambda q \frac{y_1^q + y_2^q}{q} \geq \lambda \mathbf{x} \cdot \mathbf{y}$$

Now we can at least hope to achieve the case of equality by suitably choosing  $\lambda = \lambda_0$ .



## Enter Lagrangian Multipliers

Taking the cue from there, let's apply a Lagrangian Multiplier  $\lambda$ :

$$\frac{x_1^p + x_2^p}{p} + \lambda q \frac{y_1^q + y_2^q}{q} \geq \lambda \mathbf{x} \cdot \mathbf{y}$$

Now we can at least hope to achieve the case of equality by suitably choosing  $\lambda = \lambda_0$ .

$$\lambda_0^q = \frac{x_1^p + x_2^p}{y_1^q + y_2^q}$$



## Enter Lagrangian Multipliers

---

$$\frac{x_1^p + x_2^p}{p} + \lambda^q \frac{y_1^q + y_2^q}{q} \geq \lambda \mathbf{x} \cdot \mathbf{y}$$

$$\lambda_0^q = \frac{x_1^p + x_2^p}{y_1^q + y_2^q}$$

Applying  $1/p + 1/q = 1$ , the LHS  $\frac{x_1^p + x_2^p}{p} + \lambda_0^q \frac{y_1^q + y_2^q}{q}$  simplifies to  $(x_1^p + x_2^p)$  and we have:  $(x_1^p + x_2^p) \geq \lambda_0 \mathbf{x} \cdot \mathbf{y}$ .



## Enter Lagrangian Multipliers

---

The LHS  $\frac{x_1^p + x_2^p}{p} + \lambda_0^q \frac{y_1^q + y_2^q}{q}$  simplifies to  $(x_1^p + x_2^p)$  and we have:  
 $(x_1^p + x_2^p) \geq \lambda_0 \mathbf{x} \cdot \mathbf{y}.$

One last application of  $1/p + 1/q = 1$  gives

$$(x_1^p + x_2^p)^{1/p} \cdot (y_1^q + y_2^q)^{1/q} \geq \mathbf{x} \cdot \mathbf{y}$$

and we are done!





## Enter Lagrangian Multipliers

---

The LHS  $\frac{x_1^p + x_2^p}{p} + \lambda_0^q \frac{y_1^q + y_2^q}{q}$  simplifies to  $(x_1^p + x_2^p)$  and we have:  
 $(x_1^p + x_2^p) \geq \lambda_0 \mathbf{x} \cdot \mathbf{y}.$

One last application of  $1/p + 1/q = 1$  gives

$$(x_1^p + x_2^p)^{1/p} \cdot (y_1^q + y_2^q)^{1/q} \geq \mathbf{x} \cdot \mathbf{y}$$

and we are done!

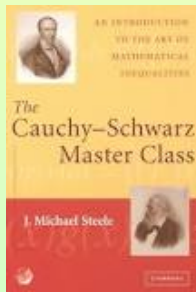
The inequality for  $n$  quantities is similar.



## *Ode to beautiful inequalities*

The world of inequalities is mesmerizing:

- ▶ The mother of all inequalities: Cauchy Schwarz Inequality  
 $(x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}(y_1^2 + y_2^2 + \cdots + y_n^2)^{1/2} \geq \mathbf{x} \cdot \mathbf{y}.$





## *Ode to beautiful inequalities*

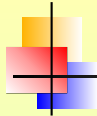
---

Sample applications:

1.  $(x^2 + y^2 + z^2) \geq (xy + yz + zx)$ ; restatement (AM-GM for 3 numbers):  $x^3 + y^3 + z^3 \geq 3xyz$ .
2. Loomis-Whitney Inequality: Compares the volume of a set in terms of the volumes of the projections of that set onto lower dimensional subspaces.

For 3 dimensions: given a set  $A \in \mathbb{R}^3$  and its projections onto the axes  $A_x, A_y, A_z$  it holds that:

$$|A| \leq |A_x|^{1/2} |A_y|^{1/2} |A_z|^{1/2}$$



## *Ode to beautiful inequalities*

---

### 1. Harker-Kasper Inequality in Crystallography



## *Ode to beautiful inequalities*

---

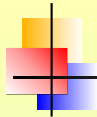
1. Harker-Kasper Inequality in Crystallography
2. Cramér-Rao Lower Bound in Statistics



## *Ode to beautiful inequalities*

---

1. Harker-Kasper Inequality in Crystallography
2. Cramér-Rao Lower Bound in Statistics
3. Jensen's Inequality in Convex Analysis, Expectation Maximization methods.



## *Ode to beautiful inequalities*

---

1. Harker-Kasper Inequality in Crystallography
2. Cramér-Rao Lower Bound in Statistics
3. Jensen's Inequality in Convex Analysis, Expectation Maximization methods.
4. Many, many, many more...



*And more...*

---

## **Carleman's Inequality:**

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} \leq e \sum_{k=1}^{\infty} a_k.$$





**THANK YOU**