2. Linearity of Expectation

- Elementary, but extremely useful tool.

- Basic fact 1: if $X_1, \ldots, X_n$ are random variables and $X = c_1 X_1 + \cdots + c_n X_n$, then:

$$E[X] = c_1 E[X_1] + \cdots + c_n E[X_n]$$

Note: Nothing needs to be assumed about the independence of the X_i.

- In applications of the technique, the X_i are often 0/1-valued indicator variables for "event i occurs", in which case for $X = X_1 + \cdots + X_n$:

$$E[X] = \text{"expected total number of events"}.$$

- Example 1. Let σ be a unif. random permutation of $[n]$ and

$$X(\sigma) = \text{number of fixed points of } \sigma.$$

Consider events "i is a fixpoint of σ" and their indicators:

$$X_i(\sigma) = \begin{cases} 1, & \text{if } \sigma(i) = i; \\ 0, & \text{if } \sigma(i) \neq i. \end{cases}$$

Clearly

$$E[X_i] = 1 \cdot \Pr(\sigma(i) = i) + 0 \cdot \Pr(\sigma(i) \neq i)$$
$$= \Pr(\sigma(i) = i) = \frac{1}{n},$$

and so

$$E[X] = E[X_1 + \cdots + X_n] = \frac{1}{n} + \cdots + \frac{1}{n} = 1.$$

- Basic fact 2: If X is a random variable on a probability space $(\Omega, \mathcal{F}, \Pr)$, then there are $\omega, \omega' \in \Omega$ s.th.

$$X(\omega) \leq E[X], \quad X(\omega') \geq E[X].$$

- Theorem 2.1 (Szele 1943).

For any n, there is a tournament T with n players and at least $2n!/2^n$ Hamiltonian paths.

[Alon (1990): Any n-player tournament contains at most $n!/(2-o(1))^n$ Hamiltonian paths.]

Proof. Let $X = X(T)$ be the number of Hamiltonian paths in a (unif.) random tournament T with n players.

For any permutation σ of $[n]$, let X_σ be the indicator variable for σ yielding a Hamiltonian path in T, i.e.

$$X_\sigma(T) = \begin{cases} 1, & \text{if in } T: \sigma(1) \to \sigma(2) \to \dots \to \sigma(n); \\ 0, & \text{otherwise.} \end{cases}$$

Then $X = \sum_\sigma X_\sigma$ and

$$\begin{aligned} E[X] &= \sum_\sigma E[X_\sigma] \\ &= \sum_\sigma \Pr(\sigma \text{ Hamiltonian w.r.t. } T) \\ &= n! \left(\tfrac{1}{2}\right)^{n-1} \\ &= 2 \cdot \frac{n!}{2^n}. \end{aligned}$$

Thus there is at least one tournament T s.th.

$$X(T) \geq E[X] = 2 \cdot \frac{n!}{2^n}. \quad \square$$

• Theorem 2.2

In any graph $G = (V, E)$, the vertices can be partitioned into $V = T \cup (V \setminus T)$, so that the number of edges crossing the cut $(T, V \setminus T)$ is at least $|E|/2$.

Proof. Consider a random cut $(T, V\setminus T)$ defined by $Pr(x \in T) = 1/2$, indep. and unif. for each $x \in V$.

Denote

$$X(T) = \text{number of edges crossing cut } T$$

and for each edge $e = \{x, y\} \in E$:

$$X_e(T) = \begin{cases} 1, & \text{if } e \text{ crosses cut } T \begin{pmatrix} x \in T,\ y \notin T \text{ or} \\ x \notin T,\ y \in T \end{pmatrix}; \\ 0, & \text{otherwise} \end{cases}$$

Then

$$E[X_e] = Pr(e \text{ crosses } T) = \frac{1}{2}$$

and

$$E[X] = E[\sum_e X_e] = \sum_e E[X_e] = |E| \cdot \frac{1}{2}.$$

Hence there is at least one cut T s.th.

$$X(T) \geq E[X] = \frac{|E|}{2}. \quad \square$$

- Now that the technique is established, the proofs of the following complicated-looking results is straightforward:

<u>Theorem</u> 2.3 For any n and k, there is a two-colouring of K_n inducing at most

$$\binom{n}{k} / 2^{\binom{k}{2}-1}$$

monochromatic K_k's. □

<u>Theorem</u> 2.4 For any m, n, h, k, there is a two-colouring of $K_{m,n}$ [the complete bipartite graph on $m+n$ vertices] inducing at most

$$\binom{m}{h}\binom{n}{k} / 2^{hk-1}$$

monochromatic $K_{h,k}$'s. □

<u>Balancing vectors</u>

- How well can linear combinations of basis vectors be approximated with simple (± 1) coefficients?

Euclidean norm

<u>Theorem</u> 2.4 Let $v_1, \ldots, v_m \in \mathbb{R}^n$, $\|v_i\| = 1 \ \forall i$. Then there exist $\varepsilon_1, \ldots, \varepsilon_m \in \{+1, -1\}$ s.th.

$$\| \varepsilon_1 v_1 + \cdots + \varepsilon_m v_m \| \leq \sqrt{m},$$

and also $\varepsilon_1, \ldots, \varepsilon_m \in \{+1, -1\}$ s.th.

$$\| \varepsilon_1 v_1 + \cdots + \varepsilon_m v_m \| \geq \sqrt{m}.$$

<u>Proof</u> Choose the $\varepsilon_i \in \{+1, -1\}$ unif. & indep. at random, and consider r.v.

$$X(\vec{\varepsilon}) = \| \varepsilon_1 v_1 + \dots + \varepsilon_m v_m \|^2$$

Then:

$$\begin{aligned} X(\vec{\varepsilon}) &= \Big(\sum_i \varepsilon_i v_i\Big)^T \Big(\sum_j \varepsilon_j v_j\Big) \\ &= \sum_i \sum_j \varepsilon_i \varepsilon_j v_i^T v_j , \end{aligned}$$

and so:

$$\begin{aligned} E[X] &= \sum_i \sum_j E[\varepsilon_i \varepsilon_j] v_i^T v_j \\ &= \sum_i \|v_i\|^2 \\ &= m. \end{aligned}$$

$$E[\varepsilon_i \varepsilon_j] = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Hence there exist specific $\vec{\varepsilon}$, $\vec{\varepsilon}'$ s.th.

$$X(\vec{\varepsilon}) \leq m, \quad X(\vec{\varepsilon}') \geq m.$$

Taking square roots completes the result. $\square$

Theorem 2.5 Let $v_1, \ldots, v_m \in \mathbb{R}^n$, $\|v_i\| \leq 1$ $\forall i$. Then for arbitrary $p_1, \ldots, p_m \in [0,1]$ and $u = p_1 v_1 + \cdots + p_m v_m$ there exist $\varepsilon_1, \ldots, \varepsilon_m \in \{0,1\}$ s.th. for $v = \varepsilon_1 v_1 + \cdots + \varepsilon_m v_m$:

$$\|u - v\| \leq \frac{\sqrt{m}}{2}.$$

Proof Idea: "approximating reals with probabilities".

Pick the $\varepsilon_i \in \{0,1\}$ indep. w/ probability $\Pr(\varepsilon_i = 1) = p_i$ and consider r.v.

$$\begin{aligned} X(\vec{\varepsilon}) &= \|u - v\|^2 \\ &= \|\sum_i (p_i - \varepsilon_i) v_i\|^2 \\ &= \sum_i \sum_j (p_i - \varepsilon_i)(p_j - \varepsilon_j) v_i^T v_j . \end{aligned}$$

Now for $i \neq j$:

$$E[(p_i - \varepsilon_i)(p_j - \varepsilon_j)] \overset{\varepsilon_i \perp \varepsilon_j}{=} E[p_i - \varepsilon_i]\, E[p_j - \varepsilon_j] = 0$$

and for $i = j$:

$$\underbrace{E[(p_i - \varepsilon_i)^2]}_{\mathrm{Var}[\varepsilon_i]} = p_i (p_i - 1)^2 + (1 - p_i) p_i^2 = p_i (1 - p_i) \leq \frac{1}{4}.$$

Thus:

$$\begin{aligned} E[X] &= \sum_i p_i (1 - p_i) \|v_i\|^2 \\ &\leq \frac{1}{4} \cdot m , \end{aligned}$$

and the rest of the proof concludes as in Thm 2.4.

□

Derandomisation

- In some (many? all?) cases a probabilistic choice can be replaced by a deterministic selection protocol; either a simple greedy one, or something more cleverly balanced.

- E.g. in the case of Thm 2.2 (large cuts), place the vertices $x \in V$ sequentially in either T or $V \setminus T$; in each case so that the majority of edges connecting x to previously considered vertices crosses the cut. Then:

$$
\begin{aligned}
X(T) &= \#\text{edges crossing } T \\
&= \sum_{e=\{x,y\}} \mathbb{1}[e=\{x,y\} \text{ crosses } T] \\
&= \sum_{x} \sum_{y<x} \mathbb{1}[e=\{x,y\} \text{ crosses } T] \\
&\geq \sum_{x} \tfrac{1}{2}\Big(\sum_{y<x} \mathbb{1}[e=\{x,y\} \in E]\Big) \\
&= \tfrac{1}{2} \sum_{x} \sum_{y<x} \mathbb{1}[e=\{x,y\} \in E] \\
&= |E|/2.
\end{aligned}
$$

- In the case of Thm 2.5 (good approximations to linear combinations), a greedy procedure also works.

Given $v_1, \dots, v_m \in \mathbb{R}^n$, $p_1, \dots, p_m \in [0,1]$, suppose $\varepsilon_1, \dots, \varepsilon_{s-1} \in \{0,1\}$ have already been chosen, and consider the choice of ε_s.

Denote by $w_t = \sum_{i=1}^{t} (p_i - \varepsilon_i) v_i$ the "error vector of the selection process by stage t.

Now given w_{s-1}, choose that value of $\varepsilon_s \in \{0,1\}$ that minimises the norm of

$$w_s = \sum_{i=1}^{s} (p_i - \varepsilon_i) v_i = w_{s-1} + (p_s - \varepsilon_s) v_s .$$

Since for <u>random</u> $\varepsilon_s \in \{0,1\}$ chosen with $\Pr(\varepsilon_s = 1) = p_s$ it holds that

$$\begin{aligned} E[\|w_s\|^2] &= \|w_{s-1}\|^2 + 2E[p_s - \varepsilon_s] w_{s-1}^T v_s \\ &\qquad + E[(p_s - \varepsilon_s)^2] \|v_s\|^2 \\ &= \|w_{s-1}\|^2 + p_s(1-p_s)\|v_s\|^2, \end{aligned}$$

there is <u>some</u> choice of ε_s that yields:

$$\|w_s\|^2 \leq \|w_{s-1}\|^2 + p_s(1-p_s)\|v_s\|^2$$

Thus, making the "greedy" choice of $\varepsilon_s \in \{0,1\}$ for all of $s = 1, \dots, m$ guarantees

$$\begin{aligned} \|w_m\|^2 &\leq \sum_{i=1}^{m} p_i(1-p_i)\|v_i\|^2 \\ &\leq \frac{1}{4} m. \end{aligned}$$