这一节课讲解了被称为独立系统的一类问题，以及用贪心解决独立系统问题的近似比。

独立系统

考虑一个有限元素集合 E，给 E 中的每个元素 e 定义一个非负的费用 $c(e)$。再考虑 $\mathcal{F} \in 2^E$，那么对于 $F \in \mathcal{F}$，我们定义 F 的费用 $c(F) = \sum\limits_{e \in F} c(e)$。现在我们要找出一个 F，使得 $c(F)$ 最大（或最小）。这就是这节课我们需要考虑的一类问题。

独立系统

从这类问题中，我们引入独立系统的概念。对于一个二元组 $(E, \mathcal{F})$，若 $\forall Y \in \mathcal{F}$，$X \subseteq Y \to X \in \mathcal{F}$，那么我们称 $(E, \mathcal{F})$ 为独立系统。由这个定义我们马上推出，$\emptyset \in \mathcal{F}$。

独立集与相关集

在独立系统 $(E, \mathcal{F})$ 中，$\mathcal{F}$ 中的元素称为独立集，$E - \mathcal{F}$ 中的元素称为相关集。

基与圈

我们将 $\mathcal{F}$ 中的极大独立集称为基，将 $E - \mathcal{F}$ 中的极小相关集称为圈。

对于 $X \subseteq E$，定义 X 上的基为 X 中的极大独立集。

秩商

对于 $X \subseteq E$，X 中的基大小可能不同。我们定义 X 的秩 $r(X)$ 为 X 中最大的基的大小，类似地定义 X 的下秩 $\rho(X)$ 为 X 中最小的基的大小。

由此定义独立系统的秩商 $q(E, \mathcal{F}) = \min\limits_{x \subseteq E} \frac{\rho(X)}{r(X)}$。秩商是一类问题中贪心解法近似比的下界，下面会进行说明。

一类最大（小）化问题

根据独立系统的定义，我们引出一类最大（小）化问题。

最大化问题：给出一个独立系统 $(E,\mathcal{F})$，找出一个 $F \in \mathcal{F}$，使得 $c(F)$ 最大。

很显然，由于每个元素的费用都是非负的，所以 $|F|$ 越大，$c(F)$ 也越大。所以最优的 F 一定是基。

最小化问题：给出一个独立系统 $(E,\mathcal{F})$，找出一个 $F \in \mathcal{F}$，使得 F 是基，且 $c(F)$ 最小。

（如果不要求 F 是基，那么取 $F=\emptyset$ 就会让代价最小，没什么意义...）

最大化问题的实例有很多。

0-1 背包问题：E 中的元素是每个物品，$\mathcal{F}$ 中的元素是所有可以放进背包的物品集合，费用就是物品的价值。

最大权独立集：E 中的元素是点，$\mathcal{F}$ 中的元素是独立集，费用就是每个点的权值。

最长简单路径：E 中的元素是边，$\mathcal{F}$ 中的元素是所有从起点到终点的简单路径以及其子集，费用就是每条边的距离。

最大权森林：E 中的元素是边，$\mathcal{F}$ 中的元素是所有不含圈的边集，费用就是每条边的权值。

最小化问题也有很多实例。

最小生成树：E 中的元素是边，$\mathcal{F}$ 中的元素是所有不含圈的边集，费用就是每条边的权值。

最短路：E 中的元素是边，$\mathcal{F}$ 中的元素是所有从起点到终点的简单路径以及其子集，费用就是每条边的距离。

旅行商问题（TSP）：E 中的元素是边，$\mathcal{F}$ 中的元素是哈密尔顿回路及其子集，费用就是每条边的距离。

拟阵

拟阵（matroid）是一个特殊的独立系统。一个独立系统需要满足以下三个条件中的一个才被称为是拟阵（事实上以下三个条件等价）：

(1) 若 $X, Y \in \mathcal{F}$，且 $|X| > |Y|$，则 $\exists e \in X - Y$，$Y \cup \{e\} \in \mathcal{F}$；

(2) 若 $X, Y \in \mathcal{F}$，且 $|X| = |Y| + 1$，则 $\exists e \in X - Y$，$Y \cup \{e\} \in \mathcal{F}$；
(3) $\forall X \subseteq E$，X 的所有基大小相同。

接下来说明这三个条件等价。

(1) 推出 (2) 是显然的，(2) 推出 (1) 使用归纳法即可。

(1) $\rightarrow$ (3)：假设存在 $X, Y \in \mathcal{F}$，X 与 Y 都是基，且 $|X| > |Y|$。那么 $\exists e \in X - Y$，$Y \cup \{e\} \in \mathcal{F}$，说明 Y 不是基。矛盾。
(3) $\rightarrow$ (1)：假设存在 $X, Y \in \mathcal{F}$，$|X| > |Y|$，且 $\forall e \in X - Y$，$Y \cup \{e\} \notin \mathcal{F}$，那么说明 Y 是基。由于 X 是独立集，存在一个基 Z 使得 $|Z| \geq |X| > |Y|$，那么有两个基 Y 与 Z 大小不同。矛盾。

我们另外定义 $\mathcal{F}^* = \{F \subseteq E \quad | \quad \exists(E, \mathcal{F}) \text{ 的基 } B, F \cap B = \emptyset\}$。

很容易发现，$(E, \mathcal{F}^*)$ 也是独立系统。我们称 $(E, \mathcal{F})$ 与 $(E, \mathcal{F}^*)$ 互为对偶。

下面证明 $F \in \mathcal{F}^{**} \rightarrow F \in \mathcal{F}$：

首先，由 $F \in \mathcal{F}^{**}$ 可以推出 $\exists(E, \mathcal{F}^*)$ 的基 $B_1, F \cap B_1 = \emptyset$。

又可以推出 $\exists(E, \mathcal{F})$ 的基 $B_2, B_1 \cap B_2 = \emptyset$。

注意到 $B_1 \cup B_2 = E$，否则我们可以从 $E - (B_1 \cup B_2)$ 中选出一个元素加入 B_1，仍有 $B_1 \cap B_2 = \emptyset$，那 B_1 就不是基了。

既然 $B_1 \cup B_2 = E$，且 $F \cap B_1 = \emptyset$，那么只能有 $F \subseteq B_2$。根据独立系统的定义，有 $F \in \mathcal{F}$。

反过来也是成立的，证明类似就略去。

两类贪心算法

下面介绍两类贪心算法，分别用于独立系统的最大化和最小化问题。

Best in：将 E 中所有元素按费用从大到小排序，使得 $c(e_1) \geq c(e_2) \geq \ldots \geq c(e_n)$。一开始令 $F = \emptyset$，按 $e_1, e_2 \ldots, e_n$ 的顺序考虑，若 e_i 加入 F 后 F 仍是独立集那就加入。这个贪心算法用于解决最大化问题。

Worst out：将 E 中所有元素按费用从大到小排序，使得 $c(e_1) \geq c(e_2) \geq \ldots \geq c(e_n)$。一开始令 $F = E$，按 $e_1, e_2 \ldots, e_n$ 的顺序考虑，若把 e_i 从 F 中去掉后 F 还含有至少一个基那就去掉。这个贪心算法用于解决最小化问题。

接下来介绍重要的 Best in 定理：设 $G(E, \mathcal{F})$ 表示 best in 贪心得到的解，$\mathrm{OPT}(E, \mathcal{F})$ 表示最优解，则 $q(E, \mathcal{F}) \leq \frac{G(E,\mathcal{F})}{\mathrm{OPT}(E,\mathcal{F})} \leq 1$ 从这个定理可以看出，如果一个独立系统是拟阵，那么用 best in 得到的最大化问题的解一定是最优解。

下面证明 Best in 定理：

首先定义 $E_j=\{e_1,e_2,\ldots,e_n\}$，$G_n$ 是 best in 贪心选中元素的集合，O_n 是最优解选中元素的集合。令 $G_j=E_j\cap G_n$ 表示 best in 贪心在考虑 e_j 之后选择了哪些元素，$O_j=E_j\cap O_n$ 表示最优解在考虑 e_j 之后选择了哪些元素。记 $d_j=c(e_j)-c(e_{j+1})$ 以及 $d_n=c(e_n)$，那么

$$\begin{aligned}
c(G_n) &= \sum_{j=1}^{n}(|G_j|-|G_{j-1}|)c(e_j) \\
&= \sum_{j=1}^{n}|G_j|d_j \\
&\geq \textstyle\sum_{j=1}^{n}\rho(E_j)d_j && \text{（因为容易证明 } G_j \text{ 是 } E_j \text{ 的一个极大独立集）} \\
&\geq q(E,\mathcal{F})\sum_{j=1}^{n}r(E_j)d_j && \text{（根据秩商的定义）} \\
&\geq q(E,\mathcal{F})\sum_{j=1}^{n}|O_j|d_j \\
&= q(E,\mathcal{F})c(O_n)
\end{aligned}$$

这就证明了 Best in 定理。

可以举一个例子说明 Best in 定理的下界是紧的：根据秩商的定义，$\exists X\subset E$，X 的基 B_1 和 B_2 满足 $\frac{|B_1|}{|B_2|}=q(E,\mathcal{F})$。我们定义 $c(e)=\begin{cases}1 & e\in X\\ 0 & e\notin X\end{cases}$ 然后把 B_1 中的元素排在前面形成 $e_1,e_2,\ldots,e_{|B_1|}$，后面随便排。如果使用 best in 贪心，就会把前面 $|B_1|$ 个元素选走，然而最优解可以选 $|B_2|$ 个元素。

另外还有两个奇怪的定理，上课提了一下...

Worst out 定理：使用 worst out 贪心得到的解满足 $1\leq\frac{G(E,\mathcal{F})}{\mathrm{OPT}(E,\mathcal{F})}\leq\max\limits_{F\subseteq E}\quad\frac{|F|-\rho^*(F)}{|F|-r^*(F)}$ 其中 $\rho^*(F)$ 表示对偶独立系统中的下秩，$r^*(F)$ 表示对偶独立系统中的秩。

n 个拟阵的交：n 个拟阵的交，用贪心得到的解近似比为 $\frac{1}{n}$。