The Radon and Carathéodory Theorems

Alexander Polyanskii
http://polyanskii.com

Moscow Institute of Physics and Technology

Partition

Standard notation: $[n] = \{1, \ldots, n\}$.

Partition

Standard notation: $[n] = \{1, \dots, n\}$.
An ***r*-partition** of set S is the set of its non-empty r subsets such that every element of S belongs to exactly one of the subsets.

Partition

Standard notation: $[n] = \{1, \ldots, n\}$.
An **r-partition** of set S is the set of its non-empty r subsets such that every element of S belongs to exactly one of the subsets.
Sometimes we do not care about the number of parts and then we can omit the parameter r, that is, we say **'partition'** instead of 'r-partition'.

Partition

Standard notation: $[n] = \{1, \ldots, n\}$.
An **r-partition** of set S is the set of its non-empty r subsets such that every element of S belongs to exactly one of the subsets.
Sometimes we do not care about the number of parts and then we can omit the parameter r, that is, we say **'partition'** instead of 'r-partition'.

For example,

$$\big\{\{2n : n \in \mathbb{Z}\}, \{2n+1 : n \in \mathbb{Z}\}\big\} \text{ is a 2-partition of } \mathbb{Z},$$
$$\big\{\{1\}, \{2,5\}, \{3,4,6\}\big\} \text{ is a 3-partition of } [6]$$
$$\text{and is not a partition of } [7];$$
$$\big\{\{1,2\}, \{2,3\}\big\} \text{ is not a partition of } [3]$$
$$\big\{\emptyset, \{1\}, \{2\}\big\}.\text{is not a 3-partition of } [2].$$

Radon Theorem

Radon Theorem in $\mathbb{R}^2$

Let $X = \{x_1, x_2, x_3, x_4\}$ be set of four points in the plane $\mathbb{R}^2$. Then there is a partition $\{X_1, X_2\}$ of X such that

$$\operatorname{conv} X_1 \cap \operatorname{conv} X_2 \neq \emptyset.$$

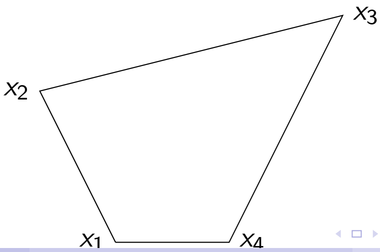
Radon Theorem in $\mathbb{R}^2$

Let $X = \{x_1, x_2, x_3, x_4\}$ be set of four points in the plane $\mathbb{R}^2$. Then there is a partition $\{X_1, X_2\}$ of X such that

$$\operatorname{conv} X_1 \cap \operatorname{conv} X_2 \neq \emptyset.$$

Proof. Consider the convex hull of X. There are several possibilities:

1. $\operatorname{conv} X$ is a quadrangle.



Radon Theorem

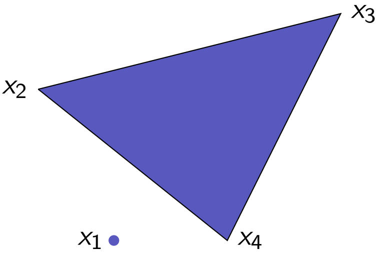
Radon Theorem in $\mathbb{R}^2$

Let $X = \{x_1, x_2, x_3, x_4\}$ be set of four points in the plane $\mathbb{R}^2$. Then there is a partition $\{X_1, X_2\}$ of X such that

$$\operatorname{conv} X_1 \cap \operatorname{conv} X_2 \neq \emptyset.$$

Proof. There are several possibilities:

1. $\operatorname{conv} X$ is a quadrangle. **Is $\{\{x_1\}, \{x_2, x_3, x_4\}\}$ a desired partition?**



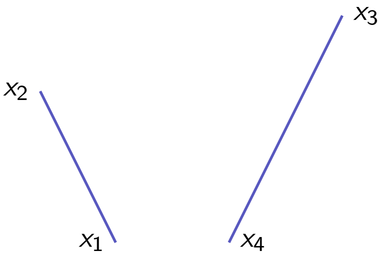
Radon Theorem in $\mathbb{R}^2$

Let $X = \{x_1, x_2, x_3, x_4\}$ be set of four points in the plane $\mathbb{R}^2$. Then there is a partition $\{X_1, X_2\}$ of X such that

$$\operatorname{conv} X_1 \cap \operatorname{conv} X_2 \neq \emptyset.$$

Proof. There are several possibilities:

1. $\operatorname{conv} X$ is a quadrangle. **Is $\{\{x_1, x_2\}, \{x_3, x_4\}\}$ a desired partition?**



Radon Theorem

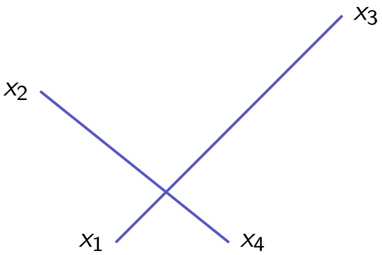
Radon Theorem in $\mathbb{R}^2$

Let $X = \{x_1, x_2, x_3, x_4\}$ be set of four points in the plane $\mathbb{R}^2$. Then there is a partition $\{X_1, X_2\}$ of X such that

$$\operatorname{conv} X_1 \cap \operatorname{conv} X_2 \neq \emptyset.$$

Proof. There are several possibilities:

1. $\operatorname{conv} X$ is a quadrangle. **Is $\{\{x_1, x_3\}, \{x_2, x_4\}\}$ a desired partition?**



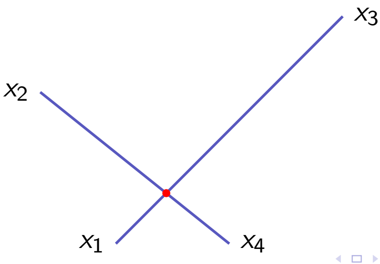
Radon Theorem in $\mathbb{R}^2$

Let $X = \{x_1, x_2, x_3, x_4\}$ be set of four points in the plane $\mathbb{R}^2$. Then there is a partition $\{X_1, X_2\}$ of X such that

$$\operatorname{conv} X_1 \cap \operatorname{conv} X_2 \neq \emptyset.$$

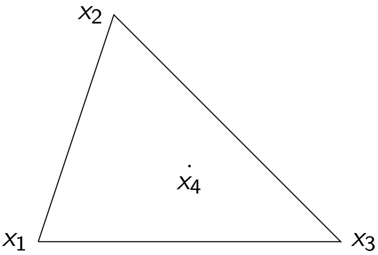
Proof. There are several possibilities:

1. $\operatorname{conv} X$ is a quadrangle. **Yes,** $\{\{x_1, x_3\}, \{x_2, x_4\}\}$ **is the desired partition!**



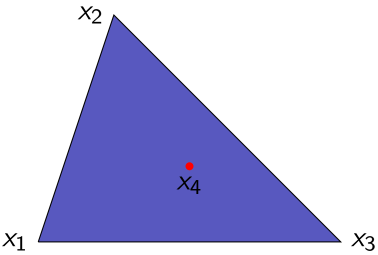
Radon Theorem: Proof

2. conv X is a triangle.



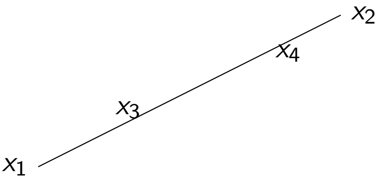
Radon Theorem: Proof

2. conv X is a triangle. $\{\{x_1, x_2, x_3\}, \{x_4\}\}$ **is the desired partition.**



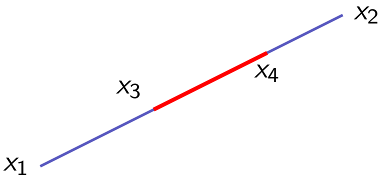
Radon Theorem: Proof

3. $\operatorname{conv} X$ is a segment.



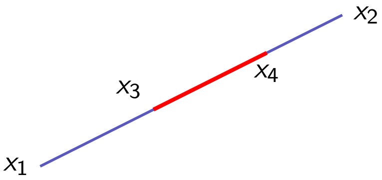
Radon Theorem: Proof

3. $\operatorname{conv} X$ is a segment. $\{\{x_1, x_2\}, \{x_3, x_4\}\}$ **is the desired partition.**



Radon Theorem: Proof

3. conv X is a segment. $\{\{x_1, x_2\}, \{x_3, x_4\}\}$ **is the desired partition.**



4. conv X is a point. $\{\{x_1, x_2\}, \{x_3, x_4\}\}$ **is the desired partition.**

$$x_1 = x_2 = x_3 = x_4$$

Radon Theorem

Radon Theorem: general case

If $X = \{x_1, x_2, \dots, x_{d+2}\}$ is a set of $d + 2$ points in $\mathbb{R}^d$. There is a 2-partition $\{X_1, X_2\}$ of X such

$$\operatorname{conv}(X_1) \cap \operatorname{conv}(X_2) \neq \emptyset.$$

Radon Theorem: general case

If $X = \{x_1, x_2, \dots, x_{d+2}\}$ is a set of $d+2$ points in $\mathbb{R}^d$. There is a 2-partition $\{X_1, X_2\}$ of X such

$$\mathrm{conv}(X_1) \cap \mathrm{conv}(X_2) \neq \emptyset.$$

Proof. $d+2$ vectors $\binom{x_i}{1}$ in $\mathbb{R}^{d+1}$ are linearly dependent, that is,

$$\binom{0}{0} = \sum_{i=1}^{d+2} \alpha_i \binom{x_i}{1}$$

for some α_i not all of them are 0's.

Radon Theorem: proof

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{d+2} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

Consider the partition $\big\{\{A_+, A_-\}\big\}$ of $[d+2]$ defined by

$$A_+ = \{i : \alpha_i > 0\} \text{ and } A_- = \{i : \alpha_i \leq 0\}.$$

Radon Theorem: proof

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{d+2} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

Consider the partition $\{\{A_+, A_-\}\}$ of $[d+2]$ defined by

$$A_+ = \{i : \alpha_i > 0\} \text{ and } A_- = \{i : \alpha_i \leq 0\}.$$

Hence,

$$\sum_{i \in A_+} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \sum_{i \in A_-} (-\alpha_i) \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

Radon Theorem: proof

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{d+2} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

Consider the partition $\{\{A_+, A_-\}\}$ of $[d+2]$ defined by

$$A_+ = \{i : \alpha_i > 0\} \text{ and } A_- = \{i : \alpha_i \leq 0\}.$$

Hence,

$$\sum_{i \in A_+} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \sum_{i \in A_-} (-\alpha_i) \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \begin{pmatrix} y \\ S \end{pmatrix},$$

where $S > 0$.

Radon Theorem: proof

$$\sum_{i \in A_+} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \sum_{i \in A_-} (-\alpha_i) \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \begin{pmatrix} y \\ S \end{pmatrix},$$

where $S > 0$. Therefore,

$$\sum_{i \in A_+} \frac{\alpha_i}{S} \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \sum_{i \in A_-} \frac{(-\alpha_i)}{S} \begin{pmatrix} x_i \\ 1 \end{pmatrix} = \begin{pmatrix} y/S \\ 1 \end{pmatrix}.$$

Radon Theorem: proof

$$\sum_{i \in A_+} \alpha_i \binom{x_i}{1} = \sum_{i \in A_-} (-\alpha_i) \binom{x_i}{1} = \binom{y}{S},$$

where $S > 0$. Therefore,

$$\sum_{i \in A_+} \frac{\alpha_i}{S} \binom{x_i}{1} = \sum_{i \in A_-} \frac{(-\alpha_i)}{S} \binom{x_i}{1} = \binom{y/S}{1}.$$

So, the point y/S belongs to $\operatorname{conv}\{x_i : i \in A_+\}$ and to $\operatorname{conv}\{x_i : i \in A_-\}$ at the same time. Thus, $\big\{\{x_i : i \in A_+\}, \{x_i : i \in A_-\}\big\}$ is the desired 2-partition.

Carathéodory Theorem

Carathéodory Theorem

If $X \subset \mathbb{R}^d$ is a finite set of points with $\mathbf{z} \in \operatorname{conv} X$, then there exists a subset $Y \subseteq X$ with $|Y| \leq d + 1$ such that $z \in \operatorname{conv} Y$.

Carathéodory Theorem

Carathéodory Theorem

If $X \subset \mathbb{R}^d$ is a finite set of points with $\mathbf{z} \in \operatorname{conv} X$, then there exists a subset $Y \subseteq X$ with $|Y| \leq d+1$ such that $z \in \operatorname{conv} Y$.

The fact that z lies in the convex hull of X can be written as

$$\binom{z}{1} = \sum_{i=1}^{n} \alpha_i \binom{x_i}{1}$$

for some nonnegative α_i.

Carathéodory Theorem

If $X \subset \mathbb{R}^d$ is a finite set of points with $\mathbf{z} \in \operatorname{conv} X$, then there exists a subset $Y \subseteq X$ with $|Y| \leq d + 1$ such that $z \in \operatorname{conv} Y$.

The fact that z lies in the convex hull of X can be written as

$$\binom{z}{1} = \sum_{i=1}^{n} \alpha_i \binom{x_i}{1}$$

for some nonnegative α_i. WLOG we assume that α_i are positive,

Carathéodory Theorem

Carathéodory Theorem

If $X \subset \mathbb{R}^d$ is a finite set of points with $\mathbf{z} \in \operatorname{conv} X$, then there exists a subset $Y \subseteq X$ with $|Y| \leq d+1$ such that $z \in \operatorname{conv} Y$.

The fact that z lies in the convex hull of X can be written as

$$\begin{pmatrix} z \\ 1 \end{pmatrix} = \sum_{i=1}^{n} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some nonnegative α_i. WLOG we assume that α_i are positive, n is the smallest positive integer for which there is such a linear combination.

Carathéodory Theorem

Carathéodory Theorem

If $X \subset \mathbb{R}^d$ is a finite set of points with $\mathbf{z} \in \operatorname{conv} X$, then there exists a subset $Y \subseteq X$ with $|Y| \leq d+1$ such that $z \in \operatorname{conv} Y$.

The fact that z lies in the convex hull of X can be written as

$$\begin{pmatrix} z \\ 1 \end{pmatrix} = \sum_{i=1}^{n} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some nonnegative α_i. WLOG we assume that α_i are positive, n is the smallest positive integer for which there is such a linear combination. Suppose $n > d+1$.

Proof of the Carathéodory Theorem

Thus these n vectors $\begin{pmatrix} x_i \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}$ are linearly dependent, that is,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{n} \beta_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some β_i not all of them are 0's.

Proof of the Carathéodory Theorem

Thus these n vectors $\begin{pmatrix} x_i \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}$ are linearly dependent, that is,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{n} \beta_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some β_i not all of them are 0's. In particular, there is at least one i with negative β_i. WLOG assume $\beta_1, \ldots, \beta_k > 0$ and $\beta_{k+1}, \ldots \beta_n \leq 0$.

Proof of the Carathéodory Theorem

Thus these n vectors $\begin{pmatrix} x_i \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}$ are linearly dependent, that is,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{n} \beta_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some β_i not all of them are 0's. In particular, there is at least one i with negative β_i. WLOG assume $\beta_1, \ldots, \beta_k > 0$ and $\beta_{k+1}, \ldots \beta_n \leq 0$. Hence, for any real t we have

$$\begin{pmatrix} z \\ 1 \end{pmatrix} = \sum_{i=1}^{n} (\alpha_i + t\beta_i) \begin{pmatrix} x_i \\ 1 \end{pmatrix}.$$

Proof of the Carathéodory Theorem

Thus these n vectors $\begin{pmatrix} x_i \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}$ are linearly dependent, that is,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{n} \beta_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some β_i not all of them are 0's. In particular, there is at least one i with negative β_i. WLOG assume $\beta_1, \dots, \beta_k > 0$ and $\beta_{k+1}, \dots \beta_n \leq 0$. Hence, for any real t we have

$$\begin{pmatrix} z \\ 1 \end{pmatrix} = \sum_{i=1}^{n} (\alpha_i + t\beta_i) \begin{pmatrix} x_i \\ 1 \end{pmatrix}.$$

For $t = 0$ the numbers $\alpha_i + t\beta_i$ are positive. For large positive t, the value $\alpha_i + t\beta_i$ becomes negative for $1 \leq i \leq k$ and the value $\alpha_i + t\beta_i$ is still positive for $k + 1 \leq i \leq n$.

Proof of the Carathéodory Theorem

Thus these n vectors $\begin{pmatrix} x_i \\ 1 \end{pmatrix} \in \mathbb{R}^{d+1}$ are linearly dependent, that is,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum_{i=1}^{n} \beta_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}$$

for some β_i not all of them are 0's. In particular, there is at least one i with negative β_i. WLOG assume $\beta_1, \ldots, \beta_k > 0$ and $\beta_{k+1}, \ldots \beta_n \leq 0$. Hence, for any real t we have

$$\begin{pmatrix} z \\ 1 \end{pmatrix} = \sum_{i=1}^{n} (\alpha_i + t\beta_i) \begin{pmatrix} x_i \\ 1 \end{pmatrix}.$$

For $t = 0$ the numbers $\alpha_i + t\beta_i$ are positive. For large positive t, the value $\alpha_i + t\beta_i$ becomes negative for $1 \leq i \leq k$ and the value $\alpha_i + t\beta_i$ is still positive for $k+1 \leq i \leq n$. Choosing $t = \min\{-\alpha_i/\beta_i : 1 \leq i \leq k\}$, we are done