I. Bárány, Lecture 1

Scribers: J. Tkadlec and T. Tkocz

In the first part of this lecture we shall recall some definitions as well as basic facts concerning convexity. Then we shall present a few classical theorems with combinatorial flavour, that is the theorems of Carathéodory, Radon and Tverberg respectively.

For two points $a, b \in \mathbb{R}^d$ we define a *segment* $[a, b]$ joining a and b as the set $[a, b] = \{\alpha a + \beta b,\ \alpha, \beta \geq 0, \alpha + \beta = 1\}$. A set C in $\mathbb{R}^d$ is called *convex* if for every two points a and b in C, the segment $[a, b]$ is also contained in C. Plainly, an intersection of convex sets is a convex set.

Given n points $a_1, \ldots, a_n$ in $\mathbb{R}^d$ and real coefficients $\alpha_1, \ldots, \alpha_n$, the point $a = \alpha_1 a_1 + \ldots + \alpha_n a_n$ is called their *positive combination* if all α_i's are nonnegative and if in addition $\alpha_1 + \ldots + \alpha_n = 1$, then the point a is called a *convex combination* (*c.c.* for short) of a_i's. If the real coefficients α_i's only satisfy the condition $\alpha_1 + \ldots + \alpha_n = 1$, then the point a is called an *affine combination* of a_i's. Note the following easy observation.

Fact 1. *If C is a convex subset of $\mathbb{R}^d$ and $a_1, \ldots, a_n \in C$, then all convex combinations of $a_1, \ldots, a_n$ belong to C.*

For a subset S in $\mathbb{R}^d$ we define

$$\operatorname{conv} S = \{\text{all convex combinations of elelements of } S\},$$
$$\operatorname{pos} S = \{\text{all positive combinations of elelements of } S\}.$$

These are clearly convex sets. Moreover, $\operatorname{conv} S$ is the smallest convex set containing S. It is called the *convex hull* of S. The set $\operatorname{pos} S$ is called the *positive hull* of S. For completeness, we also recall the definition of an *affine hull*,

$$\operatorname{aff} S = \{\text{all affine combinations of elelements of } S\}.$$

Obviously, a segment in $\mathbb{R}^d$ is a convex set. Another canonical example of a convex set is a (closed) half-space, $H = \{x \in \mathbb{R}^d,\ a \cdot x \geq \alpha\}$. The set of all positive semi-definite matrices of a fixed size is also convex. To see a nontrivial example, the interested reader is encouraged to look at the following exercise.

Exercise 1. Let f be a polynomial with complex coefficients which is not constant. Show that the roots of f' lie in the convex hull of the roots of f.

Our first theorem basically says that being in a convex hull is a *very finite* property.

Theorem 1 (Carathéodory)**.** *Let A be a subset in $\mathbb{R}^d$. Suppose that $a \in \operatorname{conv} A$. Then there exists a subset B of A with $|B| \leq d+1$ such that $a \in \operatorname{conv} B$.*

Proof. For a vector x in $\mathbb{R}^d$ and a real number α by $\begin{pmatrix} x \\ \alpha \end{pmatrix}$ we mean a vector in $\mathbb{R}^{d+1}$ with the last coordinate α. The fact that a lies in the convex hull of A can be written shortly as

$$\begin{pmatrix} a \\ 1 \end{pmatrix} = \sum_{i=1}^{n} \alpha_i \begin{pmatrix} a_i \\ 1 \end{pmatrix},$$

for some nonnegative reals α_i's (the last coordinate takes care of the condition that α_i's add up to 1). Without loss of generality we can assume that all α_i's are positive. Moreover, let n be the smallest possible for which the above holds. We want to show that $n \leq d+1$. Suppose not; then the vectors $\begin{pmatrix} a_i \\ 1 \end{pmatrix}$, $i = 1, \ldots, n$ cannot be linearly independent as they lie in a $d+1$ dimensional space. Therefore we get that

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

for some β_i's which are not all equal to zero. Hence,

$$\begin{pmatrix} a \\ 1 \end{pmatrix} = \sum_{i=1}^{n} (\alpha_i + t\beta_i) \begin{pmatrix} a_i \\ 1 \end{pmatrix}$$

for every $t \in \mathbb{R}$. Since for $t = 0$ all the coefficients $\alpha_i + t\beta_i$'s are positive, they remain positive for small t and there is a choice for t, say t_0 for which at least one of the coefficients becomes 0 with the rest remaining positive. This contradicts the minimality of n, as

$$\begin{pmatrix} a \\ 1 \end{pmatrix} = \sum_{i=1}^{n} (\alpha_i + t_0\beta_i) \begin{pmatrix} a_i \\ 1 \end{pmatrix}$$

shows that the vector $\begin{pmatrix} a \\ 1 \end{pmatrix}$ can be written as a positive combination of $\begin{pmatrix} a_i \\ 1 \end{pmatrix}$'s with fewer than n nonzero coefficients. $\square$

Geometrically speaking, Carathéodory's theorem says that a convex set A in $\mathbb{R}^d$ can be covered by simplices of A (for us, a simplex in $\mathbb{R}^d$ is a convex hull of at most $d+1$ points). With the same proof, we can also obtain an analogous result for cones (sets of all positive combinations).

Theorem 2. *Let A be a subset in $\mathbb{R}^d$. Suppose that $a \in \operatorname{pos} A$. Then there exists a subset B of A with $|B| \le d$ such that $a \in \operatorname{pos} B$.*

Sometimes, the following derivatives of Carathédory's theorem can be useful in applications.

Theorem 3. *Let A be a subset in $\mathbb{R}^d$ and $b \in \mathbb{R}^d$. Suppose that $a \in \operatorname{conv} A$. Then there exists a subset B of A with $|B| \le d$ such that $a \in \operatorname{conv}(B \cup \{b\})$.*

Theorem 4. *Let A be a subset in $\mathbb{R}^d$. Suppose that $a \in \operatorname{int}\operatorname{conv} A$. Then there exists a subset B of A with $|B| \le 2d$ such that $a \in \operatorname{int}\operatorname{conv} B$.*

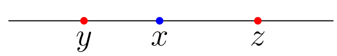
Now we move on to Radon's theorem and its further generalisation, Tverberg's theorem, saying that there are some good partitions of sets having *enough* points.

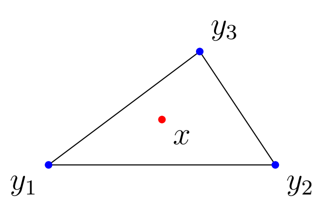
Theorem 5 (Radon)**.** *Let A be a subset in $\mathbb{R}^d$ with $|A| \ge d+2$. Then there is a partition, $A = X \cup Y$ ($X \cap Y = \varnothing$), such that $\operatorname{conv} X \cap \operatorname{conv} Y \neq \varnothing$.*

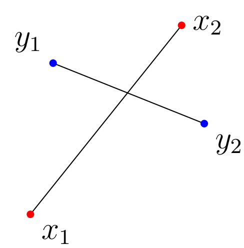
Remark. The constant $d+2$ is the best possible as shown by an example of a simplex in $\mathbb{R}^d$.

Remark. When $d = 1$ the theorem is clear as considering three points on a line, there is always one, say x between some two others, say y, z, so it suffices to take $X = \{x\}$ and $Y = A \setminus X \supset \{y, z\}$.

Remark. When $d = 2$, considering 4 points in the plane, there are two possibilities. Either certain three of them are the vertices of a triangle containing the fourth point, or the points are the vertices of a convex quadrilateral. In any case, it is clear what to take for the partition. (See the pictures.)







Proof. Since $|A| \geq d+2$, the set $\left\{ \begin{pmatrix} a \\ 1 \end{pmatrix},\ a \in A \right\}$ of vectors in $\mathbb{R}^{d+1}$ is *not* linearly independent. Therefore there are $a_i \in A$ and nonzero coefficients α_i such that

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \sum \alpha_i \begin{pmatrix} a_i \\ 1 \end{pmatrix}.$$

Because the sum of α_i's is 0, some of them are positive, some are negative. Let I be the set of all the indices i for which $\alpha_i > 0$ and J for which $\alpha_i < 0$ (neither I nor J is empty). Breaking the sum into two pieces yields

$$\sum_{i \in I} \alpha_i \begin{pmatrix} a_i \\ 1 \end{pmatrix} = \sum_{i \in J} (-\alpha_i) \begin{pmatrix} a_i \\ 1 \end{pmatrix}.$$

Dividing this by $t = \sum_{i \in I} \alpha_i = \sum_{i \in J} (-\alpha_i)$ shows that we can take $X = \{a_i,\ i \in I\}$ and $Y = A \setminus X$, as then

$$\operatorname{conv} X \ni \sum_{i \in I} \frac{\alpha_i}{t} a_i = \sum_{i \in J} \frac{-\alpha_i}{t} a_i \in \operatorname{conv} Y.$$

□

It is left as an exercise to show a combinatorial analogue of Radon's theorem. We adopt the standard notation that $[m] = \{1, 2, \ldots, m\}$.

Exercise 2. Let $A_1, \ldots, A_{n+1}$ be nonempty subsets of $[n]$. Show that there are disjoint subsets I and J of $[n+1]$ such that $\bigcap_{k \in I} A_k = \bigcap_{k \in J} A_k$. Then do the same for the union instead of the intersection.

Now we would like to present a generalisation of Radon's theorem, Tverberg's theorem, along with a sketch of a proof. There are at least seven different proofs. The one we shall sketch comes from Roudneff.

Theorem 6 (Tverberg)**.** *Let $r \geq 2$ be an integer. Let X be a subset of $\mathbb{R}^d$ with $|X| = (r-1)(d+1)+1$. Then there is a partition $X = X_1 \cup \ldots \cup X_r$ (X_i's are pairwise disjoint) such that $\bigcap_{i=1}^{r} \operatorname{conv} X_i \neq \varnothing$.*

Remark. Taking here $r = 2$ recovers Radon's theorem.

Remark. The constant $(r-1)(d+1)+1$ is the best possible. To see this, consider $(r-1)(d+1)$ points in $\mathbb{R}^d$ in general position, meaning that no $d+1$ of them lie in the same hyperplane.

Proof. We start off by stating a useful fact about intersections of affine subspaces.

Claim. *Suppose that a finite subset S of $\mathbb{R}^d$ of points lying in general position is partitioned into r pairwise disjoint subsets $S = S_1 \cup \ldots \cup S_r$ with $|S_i| \leq d+1$ for every i. If $\bigcap_{i=1}^r \operatorname{aff} S_i = \varnothing$ then $|S| \leq (r-1)(d+1)$.*

We shall not give a proof. Instead, we show a simple counting argument which we hope gives faith and indicates why the claim holds. Note that for a finite set Y of points in $\mathbb{R}^d$ in general position we have that $\operatorname{aff} Y$ is the intersection of certain $d+1-|Y|$ hyperplanes. Therefore, $\bigcap_{i=1}^r \operatorname{aff} S_i$ is the intersection of at least $d+1$ hyperplanes, as

$$\sum_{i=1}^{r} \big(d+1-|S_i|\big) = r(d+1) - |X| \geq r(d+1) - (r-1)(d+1) = d+1.$$

Hence, this intersection is empty.

Now for the proof of Tverberg's theorem, we shall consider only partitions of X into sets with at most $d+1$ elements each. Fix such a partition, $X = X_1 \cup \ldots \cup X_r$ with $|X_i| \leq d+1$, and consider the function

$$\mathbb{R}^d \ni x \overset{f}{\mapsto} \sum_{i=1}^{r} \big(\operatorname{dist}(x, \operatorname{conv} X_i)\big)^2.$$

As a sum of convex functions, it is convex, it goes to ∞ when $|x|$ does, hence it attains its minimum. Choose a partition for which this minimum is the smallest possible and suppose that the minimum equals μ and is attained at $x = z$. If $\mu = 0$ we are done, so suppose $\mu > 0$ and we will arrive at a contradiction. Choose y_i in $\operatorname{conv} X_i$ such that $|z - y_i| = \operatorname{dist}(z, \operatorname{conv} X_i)$. Now look at the function

$$\mathbb{R}^d \ni x \overset{g}{\mapsto} \sum_{i=1}^{r} |x - y_i|^2.$$

Since its value at $x = z$ is μ and it is bounded from below (pointwise) by the function f whose minimum is μ, the point $x = z$ is also where g attains its minimum. Thus, taking the gradient of g at $x = z$ yields

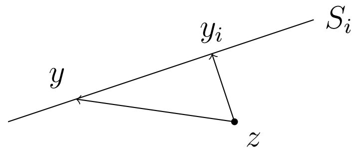
$$\sum_{i=1}^{r} (z - y_i) = 0$$

(i.e. z is the centre of mass of y_i's).

Notice that it is *not* possible that $z = y_i$ for all i's because otherwise $\mu = 0$. Without loss of generality let us assume that $z \neq y_1$, so $z \notin \operatorname{conv} X_1$. For each i, let S_i be a minimal subset of X_i for which $y_i \in \operatorname{conv} S_i$. We also have $z \notin \operatorname{conv} S_1$.

First, we show that $z \notin \cap_{i=1}^r \operatorname{aff} S_i$. If not, then in particular $z \in \operatorname{aff} S_1$. But y_1 is the point in the polytope $\operatorname{conv} S_1$ closest to z, which is outside the polytope as $z \neq y_1$, so y_1 has to lie on the boundary of the polytope which contradicts the minimality of S_1.

Second, we show that it is *not* possible for any other point $y \neq z$ to belong to $\cap_{i=1}^r \operatorname{aff} S_i$. Indeed, if not, then for every i, by the minimality of y_i,



$$(y - z) \cdot (y_i - z) \geq 0,$$

as $y \in \operatorname{aff} S_i$. Summing these inequalities yields

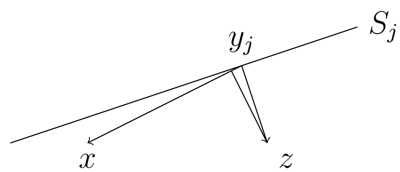
$$0 = (y - z) \cdot \sum_{i=1}^{r} (y_i - z) \geq 0,$$

and, as a result, $(y - z) \cdot (y_i - z) = 0$ for every i. Since $y \neq z$, this implies that $z = y_i$, for every i.

We have shown that $\cap_{i=1}^r \operatorname{aff} S_i = \varnothing$. By the claim we get $\sum_{i=1}^r |S_i| \leq (r-1)(d+1)$. This means that we can choose a point x in $X \setminus \bigcup_{i=1}^r S_i$. Observe that

$$\sum_{i=1}^{r} (z - y_i) \cdot (x - y_i) = \sum_{i=1}^{r} (z - y_i) \cdot \big((z - y_i) + (x - z)\big) = \sum_{i=1}^{r} (z - y_i)^2 = \mu > 0.$$

Hence, there is j for which $(z - y_j) \cdot (x - y_j) > 0$.



Moving x to the set X_j leads to a partition with a smaller minimum of f than μ (see the picture). This contradiction finishes the proof. □

Now we shall discuss Helly's theorem which also concerns intersections of convex hulls. A family of sets in $\mathbb{R}^d$ is said to have *Helly's property* if every $d+1$ of them have a nonempty intersection.

Theorem 7 (Helly). *Let $C_1, C_2, \dots, C_n$, $n \geq d+1$, be convex sets in $\mathbb{R}^d$ with Helly's property. Then*

$$\bigcap_{i=1}^{n} C_1 \neq \varnothing.$$

In other words, having empty intersection has a finite reason.

Proof. By induction on n. Case $n = d+1$ is trivial. Now suppose that $n \geq d+2$ and the theorem holds for smaller n.

For each $j = 1, \dots, n$ set

$$K_j = \bigcap_{\substack{i=1 \\ i \neq j}}^{n} C_i.$$

By induction, $K_j \neq \varnothing$. Take arbitrary $z_j \in K_j$. By Radon's Theorem, there is a partition of $z_1, \dots, z_n$ into two sets X, Y such that $\operatorname{conv} X \cap \operatorname{conv} Y \neq \varnothing$. Take z in this intersection. We claim that $z \in C_i$ for all $i = 1, \dots, n$.

Without loss of generality, focus on C_1 and suppose that $z_1 \in X$. Then all $z_j \in Y$ belong to C_1, hence $z \in C_1$ too. □

Later on, we will see another proof related to Carathéodory's Theorem.

In a special case $d = 1$ (intervals on a line), a different proof is possible. Sketch: Take the rightmost left-endpoint x of these intervals. Then every interval starts to the left of x and ends to the right of x.

Exercise 3. Extend the previous argument to trees, i.e. prove the following: If $T_1, \dots, T_n$ is a family of subtrees of a given tree T such that $T_i \cap T_j \neq \varnothing$ for each $i, j = 1, \dots, n$ then $\bigcap_{i=1}^{n} T_i \neq \varnothing$.

Note that the theorem poses no restriction on the nature of the sets C_i apart from being convex. However, it can be seen that the case with all sets compact already captures all of the complexity. Indeed, for $I \subset$

$[n+1]$, $|I| = d+1$, we can take $Z_I = \bigcap_{i \in I} C_i \neq \varnothing$ and replace C_i by $K_i = \operatorname{conv}\{Z_I,\ i \in I\} \subset C_i$.

Note that Helly's Theorem in general fails for infinite families. First example is the family of intervals of the form $I_n = [n, \infty)$, the other is the family of intervals of the form $J_n = (0, 1/n]$. However, if all the sets are compact, the theorem holds.

Theorem 8 (Helly's Theorem, infinite version)**.** *Let $C_1, C_2, \dots$ be compact convex sets in $\mathbb{R}^d$ with Helly's property. Then $\bigcap_{i=1}^{\infty} C_i \neq \varnothing$.*

Proof. Fix $n \geq 1$ and for $i = 1, \dots, n$ set $K_i = C_1 \cap \dots \cap C_i$. Then the family of the sets $K_1, \dots, K_n$ has Helly's property, hence, by usual Helly's Theorem, there exists $z_n \in \bigcap_{i=1}^{n} K_i$. By compactness, take a convergent subsequence of $\{z_n\}$ and assume it tends to some z_0. Then $z_0 \in C_i$ for all i, thus $\bigcap_{i=1}^{\infty} C_i \neq \varnothing$. □

Helly's Theorem has many applications. Here we state six of them.

1. Let $\mathcal{F}$ be a finite family of convex sets in $\mathbb{R}^d$, $|\mathcal{F}| = n \geq d+1$ and let $C \subset \mathbb{R}^d$ be convex. Then there exists a translate of C intersecting every $K \in \mathcal{F}$ if and only if there exists a translate of C intersecting every $(d+1)$-tuple from $\mathcal{F}$.

   Proof. For $K \in \mathcal{F}$, set $K^* = \{x \in \mathbb{R}^d,\ (x + C) \cap K \neq \varnothing\}$. If we prove that K^* is convex, we are done by Helly's Theorem. And indeed, take $x, y \in K^*$. Then there exist $x^* \in (x + C) \cap K \neq \varnothing$ and $y^* \in (y + C) \cap K \neq \varnothing$. Now for any $\alpha, \beta \geq 0$, $\alpha + \beta = 1$ we have

   $$\begin{aligned} &\alpha x^* \in \alpha x + \alpha C, \\ &\beta y^* \in \beta y + \beta C, \\ \text{and summing:} \qquad &\alpha x^* + \beta y^* \in \alpha x + \beta y + C. \end{aligned}$$

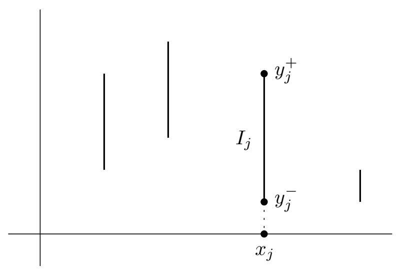
   Thus $\alpha x + \beta y + C$ is a translate of C having nonempty intersection with K, i.e. $\alpha x + \beta y \in K^*$ and K^* is convex. □

2. Let $I_1, \dots, I_n$, $n \geq 3$, be vertical intervals in $\mathbb{R}^2$ such that every three of them have a line transversal. Then all of them have a common line transversal.

Proof. A set of lines $y = \alpha x + \beta$ intersecting some particular $I_j = x_j \times [y_j^-, y_j^+]$ corresponds to a set

$$S_j = \{(\alpha, \beta) \in \mathbb{R}^2,\ y_j^- \le \alpha x_j + \beta \le y_j^+\}.$$

Being the intersections of two halfspaces, these S_j are convex, hence by Helly's Theorem they have a nonempty intersection. Any point in it determines a line intersecting all the intervals I_j.



□

3. Let $\mathcal{H}$ be a finite family of open (closed) half-spaces in $\mathbb{R}^d$ and let $C \subset \mathbb{R}^d$, $C \subset \bigcup_{H \in \mathcal{H}} H$ be convex. Then there exists a subfamily $\mathcal{H}' \subset \mathcal{H}$, $|\mathcal{H}'| = d + 1$ such that $C \subset \bigcup_{H \in \mathcal{H}'} H$.

   Proof. For all $H \in \mathcal{H}$, set $H^* = C \setminus H$. Then H^* is convex and $\bigcap_{H \in \mathcal{H}} H^* = \varnothing$. By Helly's Theorem, there exists a $(d+1)$-tuple $H_1^*, \ldots, H_{d+1}^*$ with empty intersection. Then $C \subset H_1 \cup \cdots \cup H_{d+1}$. □

4. (Kirchberger's Theorem) Sets R and B of red and blue points in $\mathbb{R}^d$ are given. Then R and B can be strictly separated by a hyperplane if and only if for all $Y \subset R \cup B$, $|Y| \le d + 2$ one can separate the sets $Y \cap R$ and $Y \cap B$. (A hyperplane h *strictly separates* sets A and B if A lies in one open half-space determined by h and B lies in the opposite closed half-space.)

   Proof. With every $r \in R$ we associate a half-space

$$C_r = \left\{ \begin{pmatrix} a \\ \alpha \end{pmatrix} \in \mathbb{R}^{d+1},\ \begin{pmatrix} a \\ \alpha \end{pmatrix} \cdot \begin{pmatrix} r \\ -1 \end{pmatrix} > 0 \right\}.$$

Likewise, with every $b \in B$ we associate

$$D_b = \left\{ \begin{pmatrix} a \\ \alpha \end{pmatrix} \in \mathbb{R}^{d+1}, \ \begin{pmatrix} a \\ \alpha \end{pmatrix} \cdot \begin{pmatrix} b \\ -1 \end{pmatrix} \leq 0 \right\}.$$

By assumption, every $d+2$ half-spaces have a point in common, hence by Helly's Theorem all the half-spaces have a point in common. This point determines a strictly separating hyperplane. □

5. (Centrepoint Theorem) Let $X \subset \mathbb{R}^d$, $|X| \geq d+1$. Then there exists a *centrepoint* of X, i.e. a point $z \in \mathbb{R}^d$ such that any closed half-space H containing z satisfies

$$|X \cap H| \geq \frac{1}{d+1}|X|.$$

Proof. Let

$$\mathcal{H} = \left\{ H, \ H \text{ is an open half-space such that } |X \cap H| < \frac{1}{d+1}|X| \right\}.$$

By summing the sizes we see that no $(d+1)$-tuple of sets in $\mathcal{H}$ covers the whole X. Hence the family of closed half-spaces satisfying $|X \cap H| \geq \frac{1}{d+1}|X|$ has Helly's property. The compactness issues in the infinite Helly's Theorem are avoided by focusing on a sufficiently large ball instead of the whole $\mathbb{R}^d$. □

6. Let $X \subset \mathbb{R}^d$ be finite with diameter $\operatorname{diam}(X) \leq 2$. Then there exists a ball B of radius $r = \sqrt{\frac{2d}{d+1}}$ that contains X. Moreover, the cases requiring $r = \sqrt{\frac{2d}{d+1}}$ are precisely the regular simplices with side length 2.

Proof. For any $x \in X$, take the ball $B(x, r)$ with centre x and radius $r = \sqrt{\frac{2d}{d+1}}$. We want to prove $\bigcap_{x \in X} B(x, r) \neq \varnothing$. By Helly's Theorem, we only need to prove it for a $(d+1)$-tuple $x_0, x_1, \ldots, x_d \in X$.

Let $B(y, R)$ be the smallest ball containing X. By appropriate translation we may assume that $y = 0$. Suppose that $|x_i| = R$ for $i = 0, 1, \ldots, m \leq d$ and $|x_i| < R$ for the rest. We will only deal with $x_0, x_1, \ldots, x_m$.

If $y \notin \operatorname{conv}\{x_0, x_1, \ldots, x_m\}$, we could move y closer, thereby reducing the radius. Assume otherwise. Then there exist $\alpha_0, \alpha_1, \ldots, \alpha_m \geq 0$ with

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \sum_{i=0}^{m} \alpha_i \begin{pmatrix} x_i \\ 1 \end{pmatrix}.$$

Now

$$4 \geq |x_i - x_j|^2 = |x_i|^2 + |x_j|^2 - 2x_i \cdot x_j = 2R^2 - 2x_i \cdot x_j.$$

Multiplying by $\alpha_i/2$ and summing over $i = 0, 1, \ldots, m$, $i \neq j$, we get

$$2(1-a_j) \geq R^2 \cdot (1-a_j) + (-\sum_{i \neq j} \alpha_i x_i) \cdot x_j = R^2 \cdot (1-a_j) + \alpha_j \cdot R^2 = R^2.$$

Summing over all $j = 0, 1, \ldots, m$ finally gives the desired

$$2(m+1-1) \geq (m+1)R^2.$$

□