To clarify, 0 in my solutions can mean the real number 0 or the zero vector or the origin point in the Euclidean space. One might easily judge the meaning from the context. I will not distinguish a point and the vector from 0 to it either.

**Q1:** (a): Let $m, n \in f(C)$ be two arbitrary points. Then we can write $m = Bx + c$ and $n = By + c$ for some $x, y \in C$. Since $C$ is convex, $\lambda x + (1-\lambda)y$ lies in $C$ for all $\lambda \in [0,1]$. Then

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= B(\lambda x + (1-\lambda)y) + c \\ &= \lambda Bx + \lambda c + (1-\lambda)By + (1-\lambda)c \\ &= \lambda(Bx + c) + (1-\lambda)(By + c) \\ &= \lambda m + (1-\lambda)n \end{aligned}$$

This means $\lambda m + (1-\lambda)n$ lies in $f(C)$ for all $\lambda \in [0,1]$. Since $m, n$ are arbitrary, $f(C)$ is convex.

(b): Not exactly. Let $f$ be the zero map, namely, $f(x) = 0$ for all $x$. Then no matter what the shape $A$ is of, $f(A) = \{0\}$ is convex. Surely, one can take $A$ to be non-convex. But the answer is positive if the matrix $B$ is invertible (which means $d = k$) since then we can find an inverse affine transformation of $f$.

(c): From part (a), we know that $f(\mathrm{conv}(X))$ is convex. So $\mathrm{conv}(f(X)) \subset f(\mathrm{conv}(X))$, since $f(X) \subset f(\mathrm{conv}(X))$.

To see the reverse inclusion, let $m \in f(\mathrm{conv}(X))$ be an arbitrary point and say $m = f(x)$ for some $x \in \mathrm{conv}(X)$. By Carathéodory's theorem, we can write $x = \sum_{i=1}^{d+1} \lambda_i x_i$, where $x_i$'s are points in $X$ (possibly with repetitions) and all $\lambda_i \geq 0$ and $\sum_{i=1}^{d+1} \lambda_i = 1$. Observe that

$$\begin{aligned} m = f(x) &= f(\sum_{i=1}^{d+1} \lambda_i x_i) \\ &= B(\sum_{i=1}^{d+1} \lambda_i x_i) + c = B(\sum_{i=1}^{d+1} \lambda_i x_i) + \sum_{i=1}^{d+1} \lambda_i c \\ &= \sum_{i=1}^{d+1} \lambda_i (Bx_i + c) = \sum_{i=1}^{d+1} \lambda_i f(x_i) \in \mathrm{conv}(f(X)). \end{aligned}$$

So indeed we have $f(\mathrm{conv}(X)) \subset \mathrm{conv}(f(X))$. And so $f(\mathrm{conv}(X)) = \mathrm{conv}(f(X))$.

**Q2:** It is clear that $\mathrm{diam}(X) \le \mathrm{diam}(\mathrm{conv}(X))$, since $X \subset \mathrm{conv}(X)$.

The case that $|X|$ is finite is easy. Say $X = \{P_1, \ldots, P_n\}$. Then an arbitrary point $x$ in $\mathrm{conv}(X)$ can be expressed as $x = \sum_{i=1}^n \lambda_i P_i$ where $\lambda_i \ge 0$ and $\sum_{i=1}^n \lambda_i = 1$. Hence

$$\begin{aligned}
\|x - P_j\| &= \|\sum_{i=1}^n \lambda_i P_i - \sum_{i=1}^n \lambda_i P_j\| \\
&= \|\sum_{i=1}^n \lambda_i (P_i - P_j)\| \\
&\le \sum_{i=1}^n \lambda_i \|P_i - P_j\| \\
&\le \sum_{i=1}^n \lambda_i \sup\{\|P_i - P_j\| : i, j = 1, \ldots, n\} \\
&= \sum_{i=1}^n \lambda_i \mathrm{diam}(X) = \mathrm{diam}(X).
\end{aligned}$$

Now for arbitrary two points $x, y \in \mathrm{conv}(X)$, we have

$$\begin{aligned}
\|y - x\| &= \|\sum_{i=1} \lambda_i y - \sum_{i=1} P_i\| \\
&\le \sum_{i=1}^n \lambda_i \|y - P_i\| \\
&\le \sum_{i=1}^n \lambda_i \sup\{\|y - P_j\| : j = 1, \ldots, n\} \\
&= \sum_{i=1}^n \lambda_i \mathrm{diam}(X) = \mathrm{diam}(X),
\end{aligned}$$

which implies $\mathrm{diam}(\mathrm{conv}(X)) \le \mathrm{diam}(X)$. Therefore, we have $\mathrm{diam}(\mathrm{conv}(X)) = \mathrm{diam}(X)$. Note that $X$ is finite set and so the equality can be achieved, which is the case $x = P_i$ and $y = P_j$ with $|P_i - P_j| = \mathrm{diam}(X)$.

Although the question says we can assume $X$ is finite, the conclusion is still true if $|X|$ is infinite. The only extra piece we need is the Carathéodory's theorem. Say $x$ is a convex combination of $P_1, \ldots, P_{d+1}$ and $y$ is a convex combination of $Q_1, \ldots, Q_{d+1}$, then the similar inequalities as above give us

$$\|x - y\| \le \sup\{||P_i - Q_j|| : i, j = 1, \ldots, d+1\} \le \mathrm{diam}(X),$$

and so $\mathrm{diam}(\mathrm{conv}(X)) \le \mathrm{diam}(X)$. We reach the same conclusion $\mathrm{diam}(\mathrm{conv}(X)) = \mathrm{diam}(X)$.

**Q3:** (a): The case $|X| = 0$ is trivial and the case $|X| = 1$ follows from Helly's theorem immediately.

For the most general case, we mimic the proof of Helly's theorem. We do by the induction on $n$. The case $n = d + 1$ is trivial. Now let $C_1, \dots, C_n$ be such sets with $n \geq d + 2$. Let $D_i = \cap_{j \neq i} C_j$. Then by induction hypothesis, each $D_i$ contains a translation of $K$. Namely, we can find $x_i$ such that $x_i + k \in D_i$ for every $k \in K$, in short $x_i + K \subset D_i$, for each $i$. Now by Radon's lemma, there are disjoint index sets $I_1, I_2 \subset \{1, 2, \dots, n\}$ such that

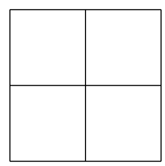
$$\operatorname{conv}(\{x_i : i \in I_1\}) \cap \operatorname{conv}(\{x_i : i \in I_2\}) \neq \emptyset.$$

Pick a point $x$ in this intersection. We claim that $x + K \subset \cap_{i=1}^n C_i$. For each $C_i$, either $i \in I_1$ or $i \in I_2$. Say $i \in I_1$. Then for any $j \in I_2$ and any $k \in K$ we have

$$x_j + k \subset D_j = \cap_{k \neq j} C_k \subset C_i.$$

But $x$ lies in $\operatorname{conv}(\{x_i : i \in I_2\})$ and so $x + k$ lies in $\operatorname{conv}(\{x_i + k : i \in I_2\}) \subset C_i$. Since the choice of $k$ is arbitrary, we have that $x + K \subset C_i$, which is true for all $i$. So indeed, $x + K \subset \cap_{i=1}^n C_i$.

(b): Consider the convex sets $A = [0, 1] \times [-1, 1], B = [-1, 0] \times [-1, 1], C = [-1, 1] \times [0, 1]$ and $D = [-1, 1] \times [-1, 0]$. Geometrically, they are the rectangle coverings of the square $[-1, 1] \times [-1, 1]$.



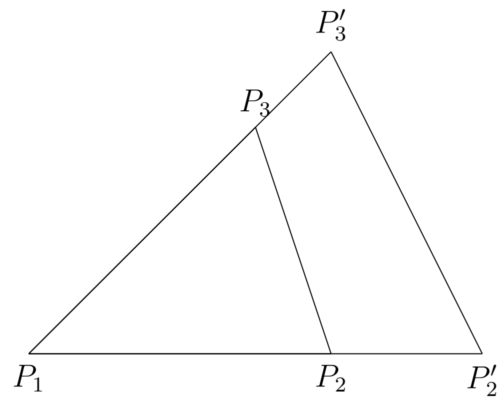
Then we can check they satisfy the assumptions that the intersection of any three of them contains a line segment of length 1. For example, $A \cap B \cap C = \{0\} \times [0, 1]$ — I believe I do not have to write all of them down. But the intersection of all of them $A \cap B \cap C \cap D = \{(0, 0)\}$. It does not contradict part (a), since the the intersections are not translations of each other but rotations.

**Q4:** (a): Say $X = \{P_1, P_2, P_3\}$. If $P_1, P_2, P_3$ are colinear, we can assume $P_3$ lies on the line segment $P_1P_2$. Let $P$ be the center of the line segment $P_1P_2$ and be a disk $D$ centered at $P$ of radius $r = \frac{1}{2}||P_1 - P_2|| = \frac{1}{2}\mathrm{diam}(X) \leq \frac{1}{2} \leq \frac{1}{\sqrt{3}}$. Then $D$ covers both ends points hence the line segment and $P_3$.

Now assume $P_1, P_2, P_3$ form a triangle. Let $P$ be the circumcenter of the triangle $\triangle P_1P_2P_3$. Then the line segments connecting $P$ to each vertex share the same length $r$, called the circumradius. The circumscribed circle (actually the disk), which centered at $P$ with radius $r$, covers $X$. Actually, we have a formula for $r$, cf Theorem 1.7.14 in the course book:

$$r = r(a,b,c) = \frac{abc}{4\sqrt{s(s-a)(s-b)(s-c)}} = \frac{abc}{\sqrt{(a+b+c)(a+b-c)(a+c-b)(b+c-a)}},$$

where $a = ||P_1P_2||, b = ||P_1P_3||, c = ||P_2P_3||$ are the three edges of the triangle and $s = \frac{1}{2}(a+b+c)$. Let us extend the line segments $P_1P_2$ (resp. $P_1P_3$) to $P_1P_2'$ (resp. $P_1P_3'$) such that $||P_1P_2'|| = 1$ (resp. $||P_1P_3'|| = 1$) — it can be the case $P_2 = P_2'$ or $P_3 = P_3'$, but it does not matter.




Consider the circumscribed disk $D'$ of $\triangle P_1P_2'P_3'$. $D'$ covers the end points $P_1, P_2', P_3'$ and also the three edges, but by construction $P_2$ (resp. $P_3$) lies in between $P_1P_2'$ (resp. $P_1P_3'$) and so is also covered by $D'$. Now the radius of $D'$ is

$$r(1,1,c) = \frac{1}{\sqrt{4-c^2}} \leq \frac{1}{\sqrt{3}}.$$

So $D'$ is the desired disk.

(b): The case where $|X| = 1$ is trivial and the case $|X] = 3$ is covered in part (a).

If $X = \{P_1, P_2\}$, then we take $P$ to be the center of the line segment $P_1P_2$ and a disk $D$ centered at $P$ of radius $l = \frac{1}{2}||P_1 - P_2|| = \frac{1}{2}\mathrm{diam}(X) \leq \frac{1}{2} \leq \frac{1}{\sqrt{3}}$. Then $D$ covers both points.

Now assume $X = \{P_1, \ldots, P_m\}$ with $m \geq 4$. For any point $P_i$, consider the disk $D_i$ of radius $\frac{1}{\sqrt{3}}$ centered at $P_i$. We claim any three disks $D_i$ must intersect non-trivially. To see this, from part (a), we know that these three points, namely, the centers of the disks, is covered by a disk of radius $\frac{1}{\sqrt{3}}$. Now the center of the covering disk lies in each disk $D_i$ and hence lies in their intersection as well. Now Helly's theorem says we can find a point $x \in \cap_{i=1}^{d} D_i$. Then the disk centered at $x$ with radius $\frac{1}{\sqrt{3}}$, which contains all the centers of $D_i$, namely, $P_i$, is the desired disk.

**Q5:** Say $A = \{P_1, \dots, P_{d+2}\}$. Since we have $d+2$ points in $\mathbb{R}^d$, they must be affine dependent, namely, we can find $\lambda_i$, not all 0 with $\sum_{i=1}^{d+2} \lambda_i = 0$, such that $\sum_{i=1}^{d+2} \lambda_i P_i = 0$. Observe that,

(1) $\lambda_i \neq 0$ for all $i$, otherwise, $A - \{P_i\}$ would be affine dependent, contradicting to the general position assumption;

(2) the choice of each $\lambda_i$ is unique up to a common multiple, in the sense that, if $\lambda_i'$, not all 0, are another choice such that $\sum_{i=1}^{d+2} \lambda_i' P_i = 0$ and $\sum_{i=1}^{d+2} \lambda_i' = 0$, then we must have $\lambda_i = t\lambda_i'$ for all $i$ for some nonzero number $t$. If not, after renaming, let us assume $\lambda_1 = t\lambda_1'$ but $\lambda_2 \neq t\lambda_2'$ for some nonzero number $t$. Since $\lambda_1$ and $\lambda_1'$ are nonzero and $\sum_{i=1}^{d+2} \lambda_i = \sum_{i=1}^{d+2} \lambda_i' = 0$, we have

$$\begin{aligned} 0 &= \lambda_1 \sum_{i=1}^{d+2} \lambda_i' - \lambda_1' \sum_{i=1}^{d+2} \lambda_i \\ &= \sum_{i=2}^{d+2} (\lambda_i'\lambda_1 - \lambda_i\lambda_1'). \end{aligned}$$

Since $\lambda_1 = t\lambda_1' \neq 0$ and $0 \neq \lambda_2 \neq t\lambda_2' \neq 0$, the first term $(\lambda_2'\lambda_1 - \lambda_2\lambda_1')$ is nonzero. Now from $\sum_{i=1}^{d+2} \lambda_i P_i = \sum_{i=1}^{d+2} \lambda_i' P_i = 0$, we see that

$$0 = \lambda_1' \sum_{i=1}^{d+2} \lambda_i P_i - \lambda_1 \sum_{i=1}^{d+2} \lambda_i' P_i = \sum_{i=2}^{d+2} (\lambda_i'\lambda_1 - \lambda_i\lambda_1') P_i,$$

which implies $P_2, \dots, P_{d+2}$ are not in general position, contradicting to our assumption.

After reordering $P_i$ and $\lambda_i$, we assume $\lambda_1, \dots, \lambda_k$ are positive and $\lambda_{k+1}, \dots, \lambda_{d+2}$ are negative. Now by the uniqueness as discussed above, $k$ (or $d+2-k$) is a fixed number. Since $\sum_{i=1}^{d+2} \lambda_i = 0$, we have $\sum_{i=1}^{k} \lambda_i = -\sum_{i=k+1}^{d+2} \lambda_i > 0$ and a unique expression up to a common multiple

$$0 = \sum_{i=1}^{d+2} \lambda_i P_i = \sum_{i=1}^{k} \lambda_i P_i - \sum_{i=k}^{d+2} (-\lambda_i) P_i.$$

Now let $\lambda = \sum_{i=1}^{k} \lambda_i = -\sum_{i=k+1}^{d+2} \lambda_i$. We claim

$$x = \frac{1}{\lambda} \sum_{i=1}^{k} \lambda_i P_i = -\frac{1}{\lambda} \sum_{i=k+1}^{d+2} \lambda_i P_i$$

is the desired point.

If $A_1, A_2$ are so partitioned that $\mathrm{conv}(A_1) \cap \mathrm{conv}(A_2) = \emptyset$, we are done. Now assume $\mathrm{conv}(A_1) \cap \mathrm{conv}(A_2) \neq \emptyset$ and say $A_1 = \{P_{i_1}, \dots, P_{i_l}\}$ and $A_2 = \{P_{i_{l+1}}, \dots, P_{i_{d+2}}\}$. Without loss of generality, we can also assume $P_1 \in A_1$. Take any $y \in \mathrm{conv}(A_1) \cap \mathrm{conv}(A_2)$. Then we can write

$$y = \sum_{j=1}^{l} c_{i_j} P_{i_j} = \sum_{j=l+1}^{d+2} c_{i_j} P_{i_j},$$

where all $c_{i_j}$ are nonnegative and $1 = \sum_{j=1}^{l} c_{i_j} = \sum_{j=l+1}^{d+2} c_{i_j}$. But then

$$0 = \sum_{j=1}^{l} c_{i_j} P_{i_j} - \sum_{j=l+1}^{d+2} c_{i_j} P_{i_j},$$

which gives the affine dependency of those $d+2$ points. By the uniqueness of the expression, we must have $l = k$ (or $l = d+2-k$), and after reordering $P_{i_j}$, we must have $P_j = P_{i_j}$ and $c_{i_j} = t\lambda_j$ for some nonzero number $t$. The number $t$ is so determined that $1 = \sum_{j=1}^k c_{i_j} = \sum_{j=1}^k t\lambda_j$ and we get $t = \frac{1}{\sum_{j=1}^k \lambda_j} = \frac{1}{\lambda}$. Then

$$y = \sum_{j=1}^k c_{i_j} P_j = \sum_{j=1}^k \frac{\lambda_j}{\lambda} P_j = x.$$

Since $y$ is arbitrary, we have $\mathrm{conv}(A_1) \cap \mathrm{conv}(A_2) = \{x\}$.

**Q6:** Say we write $R_i = [a_i, b_i] \times [c_i, d_i]$ for each rectangle $R_i$. Since each pair of rectangles have non-empty intersection, this just means for each $i, j$ we have $[a_i, b_i] \cap [a_j, b_j] \neq \emptyset$ and $[c_i, d_i] \cap [c_j, d_j] \neq \emptyset$. Now in $\mathbb{R}^1$, we apply Helly's theorem to the collection $\{[a_i, b_i] : i = 1, \ldots, k\}$ and also to the collection $\{[c_i, d_i] : i = 1, \ldots, k\}$ and we conclude that $\cap_{i=1}^k [a_i, b_i] \neq \emptyset$ and $\cap_{i=1}^k [c_i, d_i] \neq \emptyset$. And so

$$\emptyset \neq (\cap_{i=1}^k [a_i, b_i]) \times (\cap_{i=1}^k [c_i, d_i]) \subset \cap_{i=1}^k R_i.$$