Algorithmic Data Science
2nd Assignment

Nikos Stamatis
MSc Student
Graduate Programme in Data Science and Machine Learning
SN: 03400115
nikolaosstamatis@mail.ntua.gr

Exercise 1

a) Each time an update (i,Δ) arrives, the algorithm computes $h(i)$ in $O(1)$ time, $\sigma(i)$ in $O(1)$ time and the quantity $C_b+\sigma(i)\Delta$ also in $O(1)$ time. In total it requires $O(1)$ time to update C_b.

b) We define C as

$$C=\sum_{b\in[c/\varepsilon^2]} C_b^2. \tag{1}$$

The estimator C is an unbiased estimator of $\|x\|^2$:

$$\begin{aligned}\mathbb{E}[C_b^2] &= \mathbb{E}\left[\sum_{h(i)=b}\sigma(i)^2x_i^2+\sum_{\substack{h(i)=b,\\ i\neq j}}\sum_{\substack{h(j)=b,\\ j\neq i}}\sigma(i)\sigma(j)x_ix_j\right]\\ &= \mathbb{E}\left[\sum_{h(i)=b}\sigma(i)^2x_i^2\right]+\underbrace{\mathbb{E}\left[\sum_{\substack{h(i)=b,\\ i\neq j}}\sum_{\substack{h(j)=b,\\ j\neq i}}\sigma(i)\sigma(j)x_ix_j\right]}_{0}\\ &= \sum_{h(i)=b}x_i^2,\end{aligned}$$

due to the independence of $(\sigma(i))_i$. So,

$$\mathbb{E}[C]=\sum_{b\in[c/\varepsilon^2]}\mathbb{E}[C_b^2]=\sum_{b\in[c/\varepsilon^2]}\sum_{h(i)=b}x_i^2=\|x\|^2.$$

c) For computing the variance of C the following theorem will be useful:

Theorem 1. *Let $(X_i)_{i\in I}$ be an independent family of random variables and we partition I as $I=\bigcup_{j\in J}I_j$.*

1) *If $\mathcal{A}_j=\sigma\left(\{X_i : i\in I_j\}\right)$, then the family $(\mathcal{A}_j)_{j\in J}$ is independent.*
2) *If for every $j\in J$, the function $f_j:\mathbb{R}^{I_j}\to\mathbb{R}$ is measurable, and we define $Y_j=f_j((X_i)_{i\in I_j})$, then the family $(Y_j)_{j\in J}$ is independent.*

Proof. A proof for 1) can be found in [KN05, Proposition 13.14], and for 2) in [Χελ16, Appendix B']. □

By the previous theorem, the random variables $(C_b)_{b\in[c/\varepsilon^2]}$ are independent. Therefore,

$$\operatorname{Var}(C)=\operatorname{Var}\left(\sum_{b\in[c/\varepsilon^2]}C_b^2\right)=\sum_{b\in[c/\varepsilon^2]}\operatorname{Var}\left(C_b^2\right). \tag{2}$$

We can expand C_b^4 as

$$\begin{aligned}C_b^4 &= \left(\sum_{h(i)=b}\sigma(i)x_i\right)^4\\ &= \sum\sigma(i)^4x_i^4+4\sum\sigma(i)^3x_i^3\sigma(j)x_j+\\ &\quad+3\sum\sigma(i)^2x_i^2\sigma(j)^2x_j^2+\\ &\quad+6\sum\sigma(i)^2x_i^2\sigma(j)x_j\sigma(k)x_k+\\ &\quad+\sum\sigma(i)x_i\sigma(j)x_j\sigma(k)x_k\sigma(l)x_l,\end{aligned}$$

so

$$\mathbb{E}[C_b^4]=\sum_{h(i)=b}x_i^4+3\sum_{\substack{h(i)=b\\ i\neq j}}\sum_{\substack{h(j)=b\\ j\neq i}}x_i^2x_j^2$$

and

$$\begin{aligned}\operatorname{Var}(C_b^2) &= \sum_{h(i)=b}x_i^4+3\sum_{\substack{h(i)=b\\ i\neq j}}\sum_{\substack{h(j)=b\\ j\neq i}}x_i^2x_j^2-\left(\sum_{h(i)=b}x_i^2\right)^2\\ &= 2\sum_{\substack{h(i)=b\\ i\neq j}}\sum_{\substack{h(j)=b\\ j\neq i}}x_i^2x_j^2. \qquad (3)\end{aligned}$$

Combining (2) and (3) together, we obtain that

$$\begin{aligned}\operatorname{Var}(C) &= 2\sum_{b\in[c/\varepsilon^2]}\sum_{\substack{h(i)=b\\ i\neq j}}\sum_{\substack{h(j)=b\\ j\neq i}}x_i^2x_j^2 \qquad (4)\\ &= 2\sum_{b\in B}(\|x_b\|_2^4-\|x_b\|_4^4), \qquad (5)\end{aligned}$$

where $B=[c/\varepsilon^2]$, $A_b=\{i : h(i)=b\}$ and $x_b=xI_{A_b}$ is the vector all the coordinates of which are equal to zero, except for the indices that belong to A_b, on

which $x_b(i) = x(i)$. In order to bound the variance of C successfully, one needs to find an efficient bound for (5).

An easy application of the Holder inequality provides the following inequalities for the ℓ_p norms of a vector $x \in \mathbb{R}^n$ when $0 < p < q < \infty$:

$$\|x\|_q \le \|x\|_p \le n^{\frac{1}{p}-\frac{1}{q}} \|x\|_q. \tag{6}$$

Applying it to the ℓ_2 and ℓ_4 norms we obtain that $\|x\|_4^4 \le \|x\|_2^4 \le n\|x\|_4^4$, which implies that $-\|x\|_4^4 \le -\frac{1}{n}\|x\|_2^4$.

A very crude bound that ignores the buckets B is to apply this inequality to (4) to obtain that

$$\mathrm{Var}(C) \le 2(\|x\|_2^4 - \|x\|_4^2) \le 2\left(1-\frac{1}{n}\right)\|x\|_2^4, \tag{7}$$

and then proceed as in the AMS sketch proof, averaging m independent estimators $C_1, \ldots, C_m$ to obtain the estimator $\tilde{C} = \frac{1}{m}(C_1 + \ldots + C_m)$ with expectation $\|x\|_2^2$ and variance $V(\tilde{C}) \le 2\left(1-\frac{1}{n}\right)\frac{\|x\|_2^4}{m}$. Chebyshev's inequality applied to $\tilde{C}$ yields that:

$$P\left[|\tilde{C} - \|x\|_2^2| \ge \varepsilon \|x\|_2^2\right] \le \frac{\mathrm{Var}(\tilde{C})}{\varepsilon^2\|x\|_2^4} \le \frac{2\left(1-\frac{1}{n}\right)}{m\varepsilon^2}, \tag{8}$$

so by picking an $m \ge \frac{6\left(1-\frac{1}{n}\right)}{\varepsilon^2}$ we can assure that with probability at least 2/3 the estimator will $(1 \pm \varepsilon)$ approximate the quantity $\|x\|_2^2$.

However, this method is just the AMS sketch in disguise and completely ignores the second hash function h. The spirit of the exercise would suggest to only use a single estimator C and try to bound its variance efficiently by picking a sufficient number of buckets.

Revisiting (4), one can easily see that as $|B|$ tends to infinity, the sets A_b become singletons with high probability, thus $\lim_{|B|\to\infty} \mathrm{Var}(C) = 0$. So, in theory, we can achieve arbitrarily low variance by picking a sufficiently large bucket size $|B|$. The issue with this procedure is that as $|B|$ increases, the collision probability tends to zero, implying that each counter C_b will contain information about a single coordinate. In other words, storing $(C_b)_{b\in B}$ would be equivalent to storing the whole vector $(x_i)_{i=1}^n$, which is of course not desirable.

We continue from (5) to see if we can achieve a meaningful bound. Let $n_b = |A_b|$ and $n_* = \max_b\{n_b\}$. Then,

$$\begin{aligned}
V(C) &= 2\sum_{b\in B}(\|x_b\|_2^4 - \|x_b\|_4^2) \\
&\le 2\sum_{b\in B}\left(1-\frac{1}{n_b}\right)\|x_b\|_2^4 \\
&\le 2\left(1-\frac{1}{n_*}\right)\sum_{b\in B}\|x_b\|_2^4 \\
&= 2\left(1-\frac{1}{n_*}\right)\left[\left(\sum_{b\in B}\|x_b\|_2^2\right)^2 - \sum_{b\neq b'}\sum_{b'\neq b}\|x_b\|_2^2\|x_{b'}\|_2^2\right] \\
&= 2\left(1-\frac{1}{n_*}\right)\left(\|x\|_2^4 - \sum_{b\neq b'}\sum_{b'\neq b}\|x_b\|_2^2\|x_{b'}\|_2^2\right) \\
&\le 2\left(1-\frac{1}{n_*}\right)\|x\|_2^4.
\end{aligned}$$

The Chebyshev inequality will then imply that

$$P\left[|\tilde{C} - \|x\|_2^2| \ge \varepsilon\|x\|_2^2\right] \le \frac{2\left(1-\frac{1}{n_*}\right)}{\varepsilon^2}, \tag{9}$$

which is smaller than $\frac{1}{3}$ when $n_* \le \frac{6}{6-\varepsilon^2}$, namely when $n_b \le \frac{6}{6-\varepsilon^2}$ for all $b \in B$. For small values of ε, this bound is close to 1 and practically useless since it tells us that we should pick a sufficiently large bucket size so that no collisions occur, but as we explained before this method is inefficient due to its large space complexity.

Exercise 2

As usually, we denote by e the vector $e = (1, \ldots, 1)$ and by $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ the vector that has all its coordinates equal to zero, except for the i-th one which is equal to 1.

We first explain the idea behind our solution: The quantity $\sum_{i=1}^n (x_i - \mu)^2$ is just the square of the ℓ_2 norm of the vector $y = (x_1 - \mu, \ldots, x_n - \mu)$. We know that the AMS sketch algorithm could approximate $\|y\|_2^2$ within an error of $(1 \pm \varepsilon)\|y\|_2^2$ with probability $\frac{2}{3}$ using $O(\varepsilon^{-2}\log n)$ space, as long as we had a way to provide the vector y to it. Luckily, given an update (i, Δ) of the vector x, it is easy to describe a series of updates that will transform y into the desired new one.

Suppose that the current vector is $x = (x_1, \ldots, x_n)$ and the current y satisfies $y = x - \mu e$. Each time an update (i, Δ) arrives, the vector $x = (x_1, \ldots, x_n)$ changes to $x' = (x_1, \ldots, x_{i-1}, x_i + \Delta, x_{i+1}, \ldots, x_n)$, and the mean value of the original vector $\mu = \frac{x_1 + \ldots + x_n}{n}$ changes to $\mu' = \frac{x_1 + \ldots + x_n}{n} + \frac{\Delta}{n} = \mu + \frac{\Delta}{n}$. Therefore, the vector y must change to

$$y' = x' - \mu' e = x + \Delta e_i - \left(\mu + \frac{\Delta}{n}\right)e = y + \Delta e_i - \frac{\Delta}{n}e.$$

This suggests that every time an update (i, Δ) is made to the vector x, the vector y must see the updates $(i, \Delta - \frac{\Delta}{n})$ and $(j, -\frac{\Delta}{n})$ for $j \neq i$. Therefore, our algorithm is as follows:

1) Each time an update (i, Δ) appears, feed to the regular AMS sketch algorithm the following n-updates:

a) $\left(i, \Delta - \frac{\Delta}{n}\right)$,
b) $\left(j, -\frac{\Delta}{n}\right)$ for $j \neq i$.

Let $x(t) = (x_1(t), \ldots, x_n(t))$ denote the vector x after its t-th update and $\mu(t) = \frac{x_1(t)+\ldots+x_n(t)}{n}$ be its mean. Let also $y(t)$ denote the vectors created after the completions of the steps a) and b) of our algorithm.

Claim 2. *For every $t \in \mathbb{N}$, we have that*

$$y(t) = x(t) - \mu(t)e. \tag{10}$$

Proof. We will use induction on t. For $t = 0$ the result clearly holds. Suppose that (10) holds for the index t and let (i, Δ) be the next update. Then

$$\begin{aligned}
x(t+1) &= (x_1(t), \ldots, x_{i-1}(t), x_i(t) + \Delta, x_{i+1}(t), \ldots, x_n(t)) \\
&= x(t) + \Delta e_i, \\
\mu(t+1) &= \mu(t) + \frac{\Delta}{n}, \\
y(t+1) &= y(t) - \left(\frac{\Delta}{n}, \ldots, \frac{\Delta}{n}, \frac{\Delta}{n} - \Delta, \frac{\Delta}{n}, \ldots, \frac{\Delta}{n}\right) \\
&= y(t) - \frac{\Delta}{n}e + \Delta e_i \\
&= x(t) - \mu(t)e - \frac{\Delta}{n}e + \Delta e_i \qquad (11) \\
&= x(t+1) - \left(\mu(t) + \frac{\Delta}{n}\right)e \\
&= x(t+1) + \mu(t+1)e,
\end{aligned}$$

where in (11) we used our inductive hypothesis. □

An immediate consequence from the previous claim, is that the AMS sketch algorithm will indeed approximate the sum $\sum_{i=1}^{n}(x_i - \mu)^2$. Note that for our algorithm, no additional space is required compared to the original AMS sketch algorithm. However, we do perform a total of n updates in each step instead of just one, so the time complexity will increase.

Exercise 3

a) For any set $\{x_1, \ldots, x_N\}$ and any given partition $\mathcal{P} = \{P_1, \ldots, P_K\}$ of $[N]$, in order to minimize the expression

$$\text{cost}_{\mathcal{P}}(x_1, \ldots, x_N) = \sum_{k=1}^{K} \sum_{i \in P_k} \|x_i - y_k\|^2, \tag{12}$$

it suffices to find the y_k's that minimize each of the inner sums, $\sum_{i \in P_k} \|x_i - y_k\|^2$.

1st solution: As a first method we present a more general proof, where $\mathbb{R}^n$ has been replaced by an arbitrary inner product space $(X, \langle \cdot, \cdot \rangle)$. The proof idea comes from a well known result in probability, which states that the expectation μ of a random variable $X : (\Omega, \mathcal{A}, P) \to \mathbb{R}$ is the unique minimizer of its mean square error, $\mu = \arg\min_t \int_\Omega (X - t)^2 dP$.

Let $x_1, \ldots, x_N \in X$ and $\mathcal{P} = \{P_1, \ldots, P_K\}$ be a given partition of $[N]$. For each index $k \in \{1, \ldots, K\}$ we set $y_k = \frac{1}{|P_k|}\sum_{i \in P_k} x_i$. Let $z \in X$. Then

$$\begin{aligned}
\sum_{i \in P_k} \|x_i - z\|^2 &= \sum_{i \in P_k} \|x_i - y_k + y_k - z\|^2 \\
&= \sum_{i \in P_k} \|x_i - y_k\|^2 + \sum_{i \in P_k} \|y_k - z\|^2 + \\
&\quad + 2 \sum_{i \in P_k} \langle x_i - y_k, y_k - z \rangle \\
&= \sum_{i \in P_k} \|x_i - y_k\|^2 + |P_k| \cdot \|y_k - z\|^2 + \\
&\quad + 2 \left\langle \sum_{i \in P_k} x_i - |P_k| y_k, y_k - z \right\rangle \\
&= \sum_{i \in P_k} \|x_i - y_k\|^2 + |P_k| \cdot \|y_k - z\|^2 + \\
&\quad + 2 \cancel{\left\langle |P_k| y_k - |P_k| y_k, y_k - z \right\rangle} \\
&= \sum_{i \in P_k} \|x_i - y_k\|^2 + |P_k| \cdot \|y_k - z\|^2.
\end{aligned}$$

So, for every $z \in X$, we have that $\sum_{i \in P_k} \|x_i - z\|^2 \geq \sum_{i \in P_k} \|x_i - y_k\|^2$ with equality if and only if $z = y_k$. Thus the $y_k = \frac{1}{|P_k|}\sum_{i \in P_k} x_i$ are the unique minimizers of the desired expressions.

2nd solution: In the second method we will minimize the expression

$$\begin{aligned}
f(z) &= \sum_{i \in P_k} (x_i - z)'(x_i - z) \\
&= \sum_{i \in P_k} x_i' x_i - 2 \sum_{i \in P_k} x_i' z + |P_k| z' z
\end{aligned}$$

using calculus in $\mathbb{R}^n$. The derivative with respect to z,

$$\nabla f(z) = -2 \sum_{i \in P_k} x_i + 2|P_k| z,$$

is equal to zero for $z^* = \frac{1}{|P_k|}\sum_{i \in P_k} x_i = y_k$, with $\nabla^2 f(z) = 2N > 0$, so z^* minimizes f. The minimum point is unique due to f being strictly convex.

b) We first remind the Johnson-Lindenstrauss lemma [BLM13, Theorem 2.13]:

> **Theorem 3** (Johnson-Lindenstrauss). *Let $A = \{x_1, \ldots, x_N\}$ be a finite subset of $\mathbb{R}^n$ of cardinality N. Then, for every $\varepsilon > 0$ there exists a linear operator*

$T : \mathbb{R}^n \to \mathbb{R}^m$ *where* $m = O(\varepsilon^{-2} \log N)$ *such that*

(13)
$$(1-\varepsilon)\|x_i - x_j\|^2 \le \|Tx_i - Tx_j\|^2 \le (1+\varepsilon)\|x_i - x_j\|^2$$

for all $i, j = 1, \dots, N$.

Lets fix some finite set $A = \{x_1, \dots, x_N\} \subseteq \mathbb{R}^n$. A rather naive approach for this exercise would be to attempt to apply the JL lemma for the collection of points $B = \{x_1, \dots, x_N, \} \cup \{y_i : i \in I\}$, where the y_i's are all the centroids that could possibly be created for the various k-partitions of $[N]$ However, the cardinality of the set B is bounded from below by the quantity $\tilde{N} = N + kS(N,k)$.[1] A lower bound $L(N,k)$ for $S(N,k)$ is $L(N,k) = \frac{1}{2}(k^2+k+2)k^{N-l-1} - 1$ [RD69, Theorem 3], and clearly applying the JL lemma for that $\tilde{N}$ leads to non logarithmic bound with respect to N, thus the conclusion of the exercise will fail.

The idea behind the proof is to avoid including the centroids altogether into the finite set, and instead to work on the original $A = \{x_1, \dots, x_N\}$. According to the following lemma, the cost of any partition can be expressed using only terms that involve the pairwise differences of the x_i's:

Lemma 4. *Let X be an inner product space and $\{x_i : i \in P\}$ be a finite set of points in it. Let also $y = \frac{1}{|P|}\sum_{i\in P} x_i$ be their centroid. Then*

(14)
$$\sum_{i,j\in P} \|x_i - x_j\|^2 = 2|P| \sum_{i\in P} \|x_i - y\|^2.$$

In particular, for any partition $\mathcal{P} = \{P_1, \dots, P_K\}$,

(15)
$$\text{cost}_{\mathcal{P}}(x_1, \dots, x_N) = \sum_{k=1}^{K} \frac{1}{2|P_k|} \sum_{i,j\in P_k} \|x_i - x_j\|^2.$$

Proof. We expand the LHS of (14):

$$\begin{aligned}
\sum_{i,j\in P} \|x_i - x_j\|^2 &= \sum_{i\in P}\sum_{j\in P} \left(\|x_i\|^2 + \|x_j\|^2 - 2\langle x_i, x_j\rangle\right) \\
&= \sum_{i\in P}\left(|P|\|x_i\|^2 + \sum_{j\in P}\|x_j\|^2 - 2\langle x_i, \sum_{j\in P} x_j\rangle\right) \\
&= 2|P|\sum_{i\in P}\|x_i\|^2 - 2\Big\|\sum_{i\in P} x_i\Big\|^2, \qquad (16)
\end{aligned}$$

and the RHS of (14):

$$\begin{aligned}
2|P|\sum_{i\in P}\|x_i - y\|^2 &= 2|P|\sum_{i\in P}\Big\|x_i - \frac{1}{|P|}\sum_{j\in P} x_j\Big\|^2 \\
&= 2|P|\sum_{i\in P}\Big(\|x_i\|^2 + \frac{1}{|P|^2}\Big\|\sum_{j\in P} x_j\Big\|^2 \\
&\qquad - \frac{2}{|P|}\langle x_i, \sum_{j\in P} x_j\rangle\Big) \\
&= 2|P|\sum_{i\in P}\|x_i\|^2 + 2\Big\|\sum_{j\in P} x_j\Big\|^2 - 4\Big\|\sum_{j\in P} x_j\Big\|^2 \\
&= 2|P|\sum_{i\in P}\|x_i\|^2 - 2\Big\|\sum_{i\in P} x_i\Big\|^2, \qquad (17)
\end{aligned}$$

which is equal to (16). □

[1] By $S(N,k)$ we denote the Stirling numbers of second kind, which count the number of ways we can partition N distinct items into k identical sets such that all the sets are nonempty.

Lastly, we mention an obvious result that describes the centroid of the image of a finite set through a linear operator:

Lemma 5. *Let X, Y be two inner product spaces, $P = \{x_1, \dots, x_N\}$ be a finite subset of X and $y = \frac{1}{|P|}\sum_{i\in P} x_i$ its centroid. Then Ty is the centroid of the set $Q = \{Tx_1, \dots, Tx_N\}$.*

Proof. By question a) we know that the centroid of Q is $z = \frac{1}{|P|}\sum_{i\in P} Tx_i$. Since T is linear, $z = T\left(\frac{1}{|P|}\sum_{i\in P} x_i\right) = Ty$. □

We can now proceed with the solution of the exercise. Fix some $\{x_1, \dots, x_N\} \subseteq \mathbb{R}^n$ and $\varepsilon \in (0, 1/2)$. By the JL lemma, there exists some linear operator $T : \mathbb{R}^n \to \mathbb{R}^m$ with $m = O(\varepsilon^{-2}\log N)$, such that

(18)
$$(1-\varepsilon)\|x_i - x_j\|^2 \le \|Tx_i - Tx_j\|^2 \le (1+\varepsilon)\|x_i - x_j\|^2$$

for all $i, j = 1, \dots, N$.

Let $\mathcal{P} = \{P_1, \dots, P_K\}$ be any partition of $[N]$. Combining all the previous tools together, we have that

$$\begin{aligned}
(1-\varepsilon)\,\text{cost}_{\mathcal{P}}(x_1, \dots, x_N) &= (1-\varepsilon)\sum_{k=1}^{K}\sum_{i\in P_k}\|x_i - y_k\|^2 \\
&= (1-\varepsilon)\sum_{k=1}^{K}\frac{1}{2|P_k|}\sum_{i,j\in P_k}\|x_i - x_j\|^2 \qquad (19) \\
&\le \sum_{k=1}^{K}\frac{1}{2|P_k|}\sum_{i,j\in P_k}\|Tx_i - Tx_j\|^2 \qquad (20) \\
&= \sum_{k=1}^{K}\sum_{i\in P_k}\|Tx_i - Ty_k\|^2 \qquad (21) \\
&= \text{cost}_{\mathcal{P}}(Tx_1, \dots, Tx_N) \qquad (22) \\
&= \sum_{k=1}^{K}\frac{1}{2|P_k|}\sum_{i,j\in P_k}\|Tx_i - Tx_j\|^2
\end{aligned}$$

$$\leq \sum_{k=1}^{K} \frac{1+\varepsilon}{2|P_k|} \sum_{i,j \in P_k} \|x_i - x_j\|^2 \tag{23}$$

$$= (1+\varepsilon) \sum_{k=1}^{K} \sum_{i \in P_k} \|x_i - y_k\|^2 \tag{24}$$

$$= (1+\varepsilon)\, \mathrm{cost}_{\mathcal{P}}(x_1, \ldots, x_N),$$

where in (19), (21) and (24) we used Lemma 4, in (20) and (23) the JL-lemma, and in (22) Lemma 5.

References

[BLM13] S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence,* Oxford University Press, 2013. ISBN: 9780199535255 Cited on p. 3

[RD69] B. Rennie, A. Dobson, On stirling numbers of the second kind, *Journal of Combinatorial Theory*, **7** (2), pp. 116–121, 1969. DOI: 10.1016/S0021-9800(69)80045-1 Cited on p. 4

[KN05] Κουμουλλής, Γ., Νεγρεπόντης, Σ., *Θεωρία Μέτρου*, Εκδόσεις Συμμετρία, 2005. URL: simmetria.gr Cited on p. 1

[Χελ16] Χελιώτης, Δ., *Ένα δεύτερο μάθημα στις πιθανότητες*, Συνδ. Ελλ. Ακ. Βιβλιοθηκών, 2016. URL: 11419/2825 Cited on p. 1