Chapter 1.3 Verifying Matrix Multiplication

행렬의 원소가 0 과 1 인 행렬 A, B 가 있다고 하자.
그리고 A 와 B 행렬 곱셈 결과를 C 라고 하자.

$$AB = C$$

위의 식이 참인지 확인하고 싶다면,
1.1 장에 나온 것과 마찬가지로 좌변의 두 행렬을 곱셈하여 우변처럼 하나의 행렬로 구해주는 프로그램을 구현한다.

그렇지만, 이 방법은 1.1 장과 마찬가지로 "버그"가 발생할 가능성이 높고, 속도가 느리다. (Deterministic 알고리즘의 최고 성능은 $O(n^{2.37})$ 이다.)

따라서, 이번에도 성능이 더 좋은 Randomized 알고리즘을 사용하여 체크하는 방법을 알아보자.

먼저 $\bar{r} = (r_1, r_2, r_3 \dots r_n) \in \{0,1\}^n$ 인 vector $\bar{r}$을 랜덤하게 선택한다.
(각 원소 $r_1, r_2, r_3 \dots r_n$에 대해서 0, 1 을 랜덤하게 선택한다.)

그리고, A(B$\bar{r}$)과 C$\bar{r}$ 을 계산한다. (B$\bar{r}$을 먼저 계산해야 시간복잡도가 늘지 않는다.)
만약, A(B$\bar{r}$) ≠ C$\bar{r}$ 이라면 AB ≠ C 이다.
아니라면 AB = C 이라고 판단한다.

위의 Randomized Algorithm 은 다음과 같은 특성을 갖는다.
1) 시간복잡도가 O(n^2) 이다,
   Deterministic 알고리즘이 최대$O(n^{2.37})$이므로, Randomized 알고리즘이 Deterministic 보다 더 빠른 것을 알 수 있다.
2) 1.1 장 알고리즘과 마찬가지로 올바른 답을 주지 않는 경우가 있다.
   A(B$\bar{r}$) ≠ C$\bar{r}$ 이면, AB ≠ C 이지만, A(B$\bar{r}$) = C$\bar{r}$ 이라고 해서, 반드시 AB=C 를 만족하는 것은 아니다.

알고리즘이 AB≠C 임에도 불구하고, A(B$\bar{r}$) = C$\bar{r}$ 인 벡터 $\bar{r}$을 고를 확률은 다음과 같다.

$$\Pr(AB\bar{r} = C\bar{r}) \leq \frac{1}{2}$$

증명)
변수 D = AB-C ≠ 0 라 하자. 그렇다면, A(B$\bar{r}$) = C$\bar{r}$은 D$\bar{r}$ = 0 을 의미한다.
D 가 0 행렬이 아니기 때문에, D 에는 0 이 아닌 원소가 반드시 한 개는 존재한다.
W.L.G (without loss generality), d_{11}이 0 이 아닌 원소라고 하자.

D$\bar{r}$ = 0 이므로, 다음과 같은 식이 성립한다.

$$\sum_{j=1}^{n} d_{1j} * r_j = 0$$

위 식을 r_1에 대하여 정리하면,

$$r_1 = -\frac{\sum_{j=2}^{n} d_{1j} * r_j}{d_{11}}$$

식 $\sum_{j=2}^{n} d_{1j} * r_j = 0$이 될 확률을 구하지 말고,
r_1을 제외한 나머지 원소들이 다 정의되어 있다고 가정하자.
그렇다면, r_1이 위의 조건을 만족하게 될 확률은 $\frac{1}{2}$ 이다.

이와 같은 방식을 "*principle of deferred decisions*"이라고 한다.

위의 증명을 형식적으로 표현하기 위해서 "Law of Total Probability"을 소개한다.

Law of Total Probability)
각각 서로 관계없는(mutually disjoint) 이벤트 $E_1, E_2, E_3, \dots, E_n$가 있다고 하고,
$\cup_{i=1}^{n} E_i = \Omega$ 라고 하자. (Ω는 sample space 이다.)

$$\Pr(B) = \sum_{i=1}^{n} \Pr(B \cap E_i) = \sum_{i=1}^{n} \Pr(B \mid E_i) * \Pr(E_i)$$

이다.

증명)
간단하게 설명해서, 모든 event 들의 합은 sample space 전체이기 때문에
좌변과 중간 식은 당연하다. 또한 우변의 식 또한 conditional probability 의 기본 정의에 따라 증명된다.

위의 법칙을 이용하여 다음과 같이 $\Pr(AB\bar{r} = C\bar{r})$을 구할 수 있다.

$$\Pr(AB\bar{r} = C\bar{r})$$
$$= \sum_{(x_1,x_2,\dots,x_n)\in\{0,1\}^n} \Pr\left((AB\bar{r} = C\bar{r}) \cap \left((r_2, r_3, \dots, r_n) = (x_2, x_3, \dots, x_n)\right)\right)$$
$$= \sum_{(x_1,x_2,\dots,x_n)\in\{0,1\}^n} \Pr\left(\left(r_1 = -\frac{\sum_{j=2}^{n} d_{1j} * r_j}{d_{11}}\right) \cap \left((r_2, r_3, \dots, r_n) = (x_2, x_3, \dots, x_n)\right)\right)$$
$$= \sum_{(x_1,x_2,\dots,x_n)\in\{0,1\}^n} \left(Pr\left(r_1 = -\frac{\sum_{j=2}^{n} d_{1j} * r_j}{d_{11}}\right) * Pr\left((r_2, r_3, \dots, r_n) = (x_2, x_3, \dots, x_n)\right)\right)$$
$$\leq \sum_{(x_1,x_2,\dots,x_n)\in\{0,1\}^n} \frac{1}{2} * Pr\left((r_2, r_3, \dots, r_n) = (x_2, x_3, \dots, x_n)\right)$$
$$= \frac{1}{2}$$