# 1 Probability and Statistics

## (combinatorics)

Evidently, C(1, 0) = C(1, 1) = 1

Suppose for N = n ∈ N+, that for all 0 <= k <= n, we have

C(n, k) = n! / (k! \* (n - k)!)

Then let N = n + 1, for all 0 <= k+1 <= n, from the property of C, we have

And for k = 0, C(n, 0) = 1 still holds.

Thus for N = n + 1, the equation still holds.

Therefore, for any N∈ N+ and any 0 <= K <= N, we have

## (counting)

1. Altogether 210 cases, each with probability (1/2)10.

Among them, C(10, 4) cases fit our demand.

P(head = 4, tail = 6) = C(10, 4) \* (1/2)10

1. Altogether 210 cases, each with probability 1 / C(52, 5).

Among them, take 2 numbers, from each number take 3/2 cards.

X – Y are commutative, thus the result is multiplied by 2.

P(full house) =

## (conditional probability)

## (Bayes theorem)

## (union/intersection)

1. Max(P(A ∩ B)) = 0.3 when A ⊆ B
2. Min(P(A ∩ B)) = 0 when A ∩ B = ∅
3. Max(P(A ∪ B)) = 0.7 when A ∩ B = ∅
4. Min(P(A ∪ B)) = 0.4 when A ⊆ B

# 2 Linear Algebra

## (rank)

It is a rank-2 square matrix.

## (inverse)

## (eigenvalues/eigenvectors)

|  |  |
| --- | --- |
| eigenvalues | eigenvectors |
| 4 | 1, 2, -1 |
| 2 | 1, 0, -1 |
| 2 | 0, 1, -1 |

## (singular value decomposition)

Obviously, if

w.l.o.g Suppose M is an m-by-n (m<n) matrix, then is an m-by-n matrix with non-negative real numbers on the diagonal.

, where

Thus,

, where

, where

Therefore for any , we have

, which means

1. w.l.o.g Suppose M is an m-by-n (m<n) matrix, then is an m-by-n matrix with non-negative real numbers on the diagonal.

M is invertible, therefore rank(M) = m. Since Umxm and Vnxn are orthogonal,

rank(U) = m, rank(V) = n > m, thus rank() >= m, which suggests

Similar to the proof in (a), we have

Therefore

This indicates that

1. (PD/PSD)
2. for all x ≠ 0, , therefore is PSD
3. =>

Let *λ*, *v* be an eigenvalue-eigenvector pair of A.

Then

Given that A is PD, there exists at least one , and

Since , has to be strictly positive.

<=

Suppose there exists an eigenvalue-eigenvector pair *λ*, *v* of A s.t. *λ <=* 0

Then , which contradicts with the fact that A is PD.

This indicates that each eigenvalue of A has to be strictly positive.

1. (inner product)

4. . If apparently any is good. Otherwise,

let w.l.o.g suppose

let

Evidently,

So we just make

# 3 Calculus

## (differential and partial differential)

1. (chain rule)
2. (gradient and Hessian)

At u = 1 and v = 1,

At u = 1 and v = 1,

1. (Taylor’s expansion)
2. (optimization)

Let , since A > 0, B > 0, we have

, therefore is the minimum.

1. (vector calculus)

Let ,,

Thus for 1<=i<=d

Which yields

Henceforth