5강

Useful norms

$$||x||_1 = \sum |x_i|$$
$$||x||_2 = \sqrt{\sum_i x_i^2}$$
$$||x||_\infty = max_i[|x_i|]$$

p = 무한대? → infinite norm

||x||_1 → 각각의 절대값을 다 더한 것임

||x||_2 → 벡터의 길이를 정의

infinite norm → 각각의 절대 값 중에서 가장 큰 값을 return하는 함수이다. 왜?

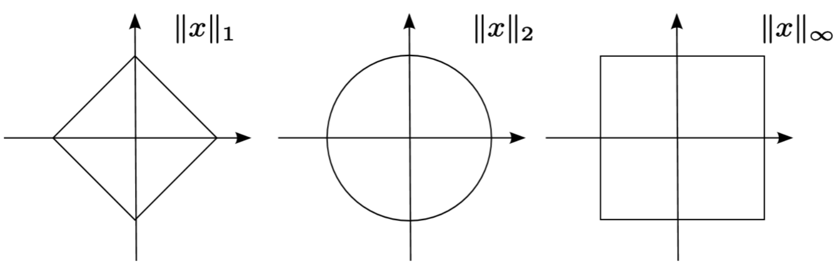
→ p가 아주 커지면 결국에는 벡터 x의 가장 큰 element만이 살아남는다.

→ 그러면 거기에 1/p를 씌우니까 그 가장 큰 element만이 살아남게 된다.

→ 이런 최댓값을 뽑아내는 개념이 중요하다!!

→ 이 함수는 미분이 불가능한 함수이다. (각이 져있다, max는 각이져있다.)

→ 그런데 우리가 max를 쓰는 것이 아니라 p값을 크게만해서 미분가능하게 만들어서 approximation해보면 안되나?



$$||x||_1 = 1 \qquad ||x||_2 = 1 \qquad ||x||_\infty = 1$$

norm도 어떤 norm을 쓰느냐에 따라 값이 달라진다.

→ p에 따라 거리가 달라지는 것이다.

전부 그래프들을 겹쳐서 생각해보라

→ 내가 똑같은 지점에 있더라도 거리가 가깝거나 멀리보이는 것이다.

Norm balls and norm cones

A (Euclidean) ball is

$$B(x_0, r) = \{x| \ ||x - x_0||_2 \leq r\}$$

→ Norm balls은 norm의 점들을 모아놓은 것이라고 생각하라

- More generally a p-norm ball with

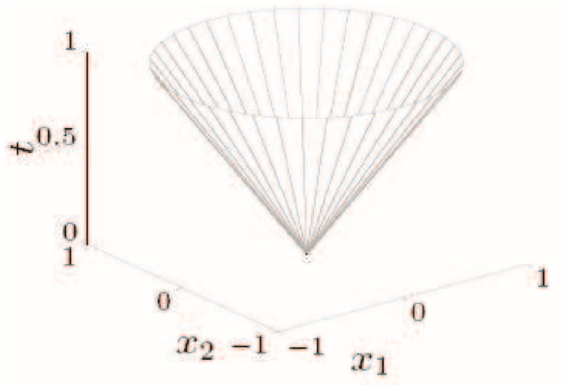
$$B_p(x_0, r) = \{x| \ ||x - x_0||_p \leq r\}$$

  Note

$$B_1(O, r) \subset B_2(O, r) \subset B_\infty(O, r)$$

- norm cone:

$$\{(x, t)|||x|| \leq t\}$$



→ x = (x_1, x_2)

→ 세로축이 t이고, 그 단면을 자른다고 생각해봐라

→ t는 0이하인 scalar이다.

→ 만일 x가 n차원이면, norm cone (x, t)는 n-1차원이다.

- Norm cones and norm balls are convex sets

  → 밖으로 볼록, 혹은 직선이기에 convex sets이다.

Inequality signs for vectors and matrices

벡터의 크기비교를 어떻게 할 것인가?

- For vectors x and y, $x \leq y$ denotes the componentwise inequality.

  → componentwise inequality은 각각 같은 차원에 있는 element끼리 비교해서 한쪽 벡터의 모든 element가 작으면 부등호 기호를 사용하는 것이다.

  여기서는 위글리, 좀 더 구부러진 부등호를 사용하기는 하는데, 수업에서는 그냥 부등호를 사용한다.

- With some abuse of notation, we will write $x \geq 0$ for componentwise nonnegative vector

  x가 벡터인데, 0보다 크다고하면, x의 모든 element가 0보다 크다는 것이다.

  → 사실 엄밀히 말하면 0벡터보다 크다는 것을 말하는 것이다.

- Can we define signs for matrices?

  행렬에 대해서 어떤 sign을 줄 수 있는가?

  → 따로 정의를 하는 개념이 있다.

Positive definite matrices

→ 행렬의 음수, 양수에 준하는 개념이다.

- A symmetric matrix A ∈ R^n×n is positive (semi-)definite

  if for any x ∈ R^n with x != 0, we have that x^T * A * x > 0 (x^T * A * x ≥ 0).

  → 일단은 nxn의 symmetric matrix을 가정한다.

  → 모든 가능한 n차원 벡터 x에 대하여(x가 0벡터가 아닌 경우에)

  다음이 성립하면 PD(Positive definite)이다.

—

$$x^T Ax > 0$$

→ 모든 가능한 n차원 벡터 x에 대하여(x가 0벡터가 아닌 경우에)

다음이 성립하면 PSD(Positive semidefinite)이다.

$$x^T Ax \geq 0$$

n=1이면, scalar이다

→ x * a * x = a * x^2

→ 0이 아닌 x에 대하여 a * x^2가 항상 양수이려면, a가 항상 양수여야 한다.

→ 즉, 이는 scalar를 포함하는 정의인 것이다.

x^T * A * x의 결과는 항상 scalar가 된다. → 0과 바로 비교 가능

n=2이면?

$$A = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$
$$x^T Ax = 1/2(x_1 + x_2)^2 + x - 1^2/2 + x_2^2/2$$

→ d=이차식이 나온다. 그 이차식이 항상 0보다 크거나 작냐를 묻는 말인 것이다.

→ PSD도 마찮가지이다.

$$A \succ 0 \ for \ PD$$
$$A \succcurlyeq 0 \ for \ PSD$$
$$A > 0 \ for \ positive \ matrix$$
$$A \geq 0 \ for \ non - negative \ matrix$$

positive matrix는 행렬의 모든 element가 다 양수라는 의미이다.

- This is a way to define “sign” of a matrix – by considering the quadratic form

  → 아래로 볼록해야 한다는 성질에 관한 것이다.

  → A가 PD냐 아니냐로 결정된다.

- PD/PSD matrices are important concepts in understanding a popular field of convex optimization – quadratic programming

→ 아주 중요한 성질이다.

- some definitions
  - S^n: set of symmetric n × n matrices
  - S^n_+: set of symmetric positive semidefinite n × n matrices (PSD)
  - S^n_++: set of symmetric positive definite n × n matrices (PD)

- PSD/PD matrices can be characterized by the sign of their eigenvalues

  → eigenvalues의 부호들을 보면 알 수 있다.
  - Real symmetric matrices have real eigenvalues

    → 허수가 아닌 실수로 봐라

$$
A = Q^T \Lambda Q, \\
\Lambda = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & \lambda_i & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \\
A \in S^n_+ \ if \ \lambda_i \geq 0 \\
A \in A^n_{++} \ if \ \lambda_i > 0
$$

  - PSD/PD implies that its eigenvalues are all nonnegative/positive

---

Polyhedra

- Solution set of finitely many linear inequalities and equalities
- A polyhedron is defined as

$$
\{x | Ax \leq b, Cx = d\} \\
for \ some \ A \in R^{m \times n}, b \in R^m, C \in R^{k \times n}, d \in R^k
$$

→ 이건 element끼리의 비교 부등호이다.

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \dots \\ a_m^T \end{bmatrix}$$

$$\begin{bmatrix} a_1^T x \\ a_2^T x \\ \dots \\ a_m^T x \end{bmatrix} \leq \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$$

constraint가 m개 인 것이다.

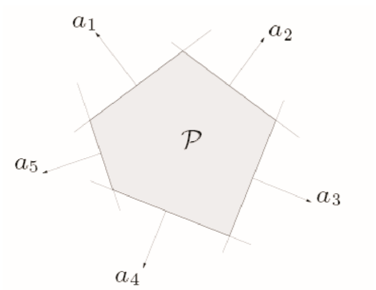
$$\begin{gathered} a_1^T \leq b_1 \\ a_2^T \leq b_2 \\ \dots \\ a_m^T \leq b_m \end{gathered}$$

→ 이는 기하학적으로 Halfspaces들의 교집합을 의미한다.

→ Cx=d는 마찮가지로 k개의 Hyperplanes들이 동시에 만나는 점들의 집합을 의미한다.

→ 즉, m개의 Halfspaces들과 k개의 Hyperplanes들의 교집합을 의미한다.

- In other words, polyhedra is the intersection of a finite number of halfspaces and hyperplanes



→ 이것은 항상 convex set이 될 수 밖에 없다.

→ 찌그러진 다각형, 다면체는 나올 수 없다.

→ convex set의 교집합들은 convex set이라는 성질이 있다.

(Halfspaces와 Hyperplanes은 모두 convex set이다.)