

## 블로그 메뉴

홈  
미디어로그  
태그  
방명록

## 검색결과 리스트

### 글

#### [선형대수학 #5] 선형연립방정식 풀이

수학 이야기 2013. 10. 29. 17:43

다양한 형태의 선형 연립 방정식을 최소자승법(least square method)과 SVD(singular value decomposition), pseudo inverse 등을 이용해서 푸는 방법을 정리해 봅니다.

먼저, 풀이 방법을 요약 정리한 후에 각각에 대해서 구체적인 예와 함께 설명합니다.

#### 1. 선형연립방정식 풀이 방법 요약

$AX = B$  형태

- A의 역행렬이 존재하는 경우:  $X = A^{-1}B$
- A의 역행렬이 존재하지 않는 경우:  $X = A^+B$  (단,  $A^+$ 는 A의 pseudo inverse)

$AX = 0$  형태

- A의 특이값분해(SVD)를  $A = U\Sigma V^T$ 라 할 때, X는 V의 가장 오른쪽 열벡터 (즉, A의 최소 특이값에 대응하는 right singular vector)

#### 2. $AX = B$ 형태

미지수의 개수가 n개, 식의 개수가 m개인 선형연립방정식은 다음과 같다 (단,  $m \geq n$ ).

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \quad \text{--- (1)}$$

위와 같은 x에 대한 선형연립방정식(a, b는 상수)을 행렬로 표현하면

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \quad \text{--- (2)}$$

$$AX=B \quad \text{--- (3)}$$

이 때,  $m = n$ 이고  $A$ 의 역행렬이 존재할 경우 위 선형방정식 (1)을 만족하는 해는 다음과 같이 유일하게 결정된다.

$$X=A^{-1}B \quad \text{--- (4)}$$

그러나, 실제로 부딪히는 대부분의 문제는  $m > n$ 인 경우로서  $A$ 의 역행렬도 식 (1)을 만족하는 해도 존재하지 않는다. 이 경우에는 다음과 같이  $A$ 의 의사역행렬(pseudo inverse)를 이용하여  $X$ 를 근사적으로 구한다.

$$X=A^+B \quad \text{--- (5)}$$

이 때,  $A$ 의 pseudo inverse  $A^+$ 는 다음과 같이 계산된다.

$$A^+=(A^T A)^{-1}A^T \quad \text{--- (6)}$$

$$A^+=V\Sigma^+U^T \quad \text{--- (7)}$$

(6), (7) 어느 방법을 사용하여 pseudo inverse를 계산해도 무방하나, 식 (6)은  $A^T A$ 의 역행렬이 존재할 경우만 적용할 수 있다 (대부분의 실제 문제들은  $A^T A$ 의 역행렬이 존재). 식 (7)은 특이값분해(SVD)를 이용해서 pseudo inverse를 계산하는 경우로서  $A$ 의 SVD가  $A = U\Sigma V^T$ 라 할 때,  $A$ 의 pseudo inverse는  $A^+ = V\Sigma^+U^T$ 가 된다 (단,  $\Sigma^+$ 는  $\Sigma$ 의 0이 아닌 원소들의 역수를 취한 후 transpose를 시킨 것).

식 (5)와 같이 pseudo inverse를 이용하여 구한  $X$ 는  $\|AX-B\|$ 를 최소화시키는  $X$ 와 동일하다. 그리고 이와같이 pseudo inverse를 이용하여 구한 근사해를 least square solution이라고 부르고, 이러한 풀이 방식을 최소자승법 (least square method)이라 부른다.

### 3. $AX = 0$ 형태

미지수의 개수가  $n$ 개, 식의 개수가  $m$ 개인  $AX = 0$  형태의 homogeneous 선형연립방정식은 다음과 같다.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \quad \text{--- (8)}$$

이를 행렬로 표현하면

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{--- (9)}$$

$$AX=0 \quad \text{--- (10)}$$

이러한 homogeneous 선형방정식의 해는 A의 최소 특이값(singular value)에 대응하는 right singular vector로 주어진다. 먼저 A의 특이값분해(SVD)가 다음과 같다고 하자.

$$A=U\Sigma V^T \quad \text{--- (11)}$$

이 때,  $\Sigma$ 의 대각선상에 위치한 원소들이 A의 특이값(singular)이고, U, V는 모두 직교행렬(orthogonal matrix), 특이값들은 모두 0 이상(0 또는 양수)임은 앞서 설명한 바 있다. 이제 식 (10)의 해는 A의 특이값에 0이 포함되는지 여부에 따라 다음과 같이 두 가지 경우로 나뉜다.

i) 특이값에 0이 포함된 경우

만일 A의 특이값들 중 0이 존재하면 0인 특이값에 대응하는 right singular vector (V의 열벡터)는 식 (10)을 만족하는 0이 아닌 해가 된다. 만일, 0의 값을 갖는 특이값이 여러개 존재할 경우에는 대응되는 right singular vector들이 모두 식 (10)의 해가 된다. 또한 이들 해의 임의의 일차결합도 역시 해가 된다.

ii) 특이값이 모두 양수인 경우

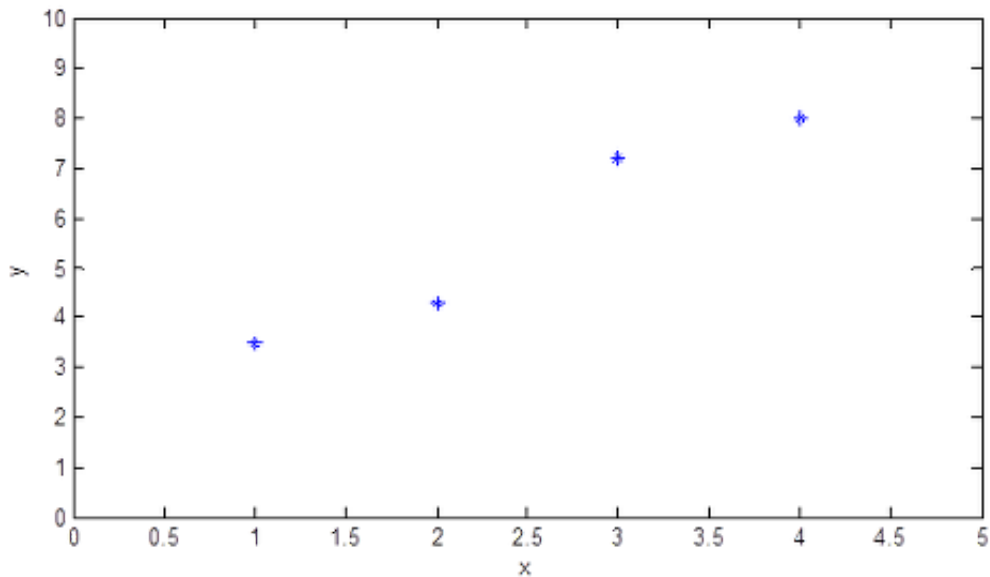
특이값이 모두 양수인 경우에는 식 (10)을 정확히 만족하는 해는  $X = 0$ 을 제외하고는 존재치 않는다. 따라서, 이 경우에는 근사적으로 해를 구해야 하는데  $\|AX\|$ 를 최소로 하는 근사해는 A의 특이값들 중 **최소 특이값에 대응하는 right singular vector(즉, V의 가장 오른쪽 열벡터)**이다.

☞ right singular vector가 해가 되는 이유를 간략히 설명해 보면, 예를 들어 i번째 singular value가 0 ( $\sigma_i = 0$ ), 대응되는 right singular vector를  $v_i$ 라 하자( $v_i$ 는 V의 i번째 열벡터). 이 때,  $Av_i = U\Sigma V^T v_i$ 인데,  $V^T v_i$ 는 i번째 값만 1이고 나머지는 0인 n차원 열벡터가 된다( $\because V$ 는 직교행렬). 그런데, 여기에  $\Sigma$ 를 곱하면  $\Sigma V^T v_i$ 은 m차원 0벡터가 된다( $\because \sigma_i = 0$ ). 따라서,  $Av_i = U\Sigma V^T v_i = 0$ 이 되므로  $v_i$ 는  $AX = 0$ 의 해가 된다.

☞ 이를 좀더 일반적으로 보면, i번째 특이값을  $\sigma_i$ , 대응되는 right singular vector를  $v_i$ , left singular vector  $u_i$ 라 했을 때 X에  $v_i$ 를 대입하면 항상  $Av_i = \sigma_i * u_i$ 가 나온다. 따라서,  $\sigma_i = 0$  이라면  $v_i$ 는  $AX = 0$ 의 해가 되고, 모든 특이값이 양수인 경우에는 최소 특이값에 대응하는  $v_i$ 가  $\|AX\| = \|Av_i\| = \|\sigma_i * u_i\| = \sigma_i$  ( $\because U$ 는 직교행렬이므로  $u_i$ 는 단위벡터)를 최소화하는 근사해가 된다. 보다 엄밀한 증명을 위해서는 임의의 벡터 X에 대해  $\|AX\| \geq \|Av_i\|$  임을 증명해야 하나 이는 내용상 중요치 않으므로 생략한다 ( $v_i$ 들이 기저를 이루기 때문에 임의의 벡터는  $v_i$ 들의 일차결합으로 표현될 수 있음을 이용하면 어렵지 않게 증명할 수 있다).

#### 4. 직선 근사 예제

예를 들어 네 점 (1, 3.5), (2, 4.3), (3, 7.2), (4, 8)을 가장 잘 근사하는 직선의 방정식을 구하는 문제를 생각해 보자.



i) 직선의 방정식을  $y = ax + b$ 로 잡은 경우

$$\begin{aligned} 1a+b &= 3.5 \\ 2a+b &= 4.3 \\ 3a+b &= 7.2 \\ 4a+b &= 8.0 \end{aligned} \quad \text{--- (12)}$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3.5 \\ 4.3 \\ 7.2 \\ 8.0 \end{bmatrix} \quad \text{--- (13)}$$

A의 pseudo inverse를 구하여 a, b를 구해보면 직선의 방정식은 다음과 같이 구해진다.

$$y = 1.64x + 1.65 \quad \text{--- (14)}$$

ii) 직선의 방정식을  $ax + by + c = 0$ 으로 잡은 경우

$$\begin{aligned} a+3.5b+c &= 0 \\ 2a+4.3b+c &= 0 \\ 3a+7.2b+c &= 0 \\ 4a+8.0b+c &= 0 \end{aligned} \quad \text{--- (15)}$$

$$\begin{bmatrix} 1 & 3.5 & 1 \\ 2 & 4.3 & 1 \\ 3 & 7.2 & 1 \\ 4 & 8.0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{--- (16)}$$

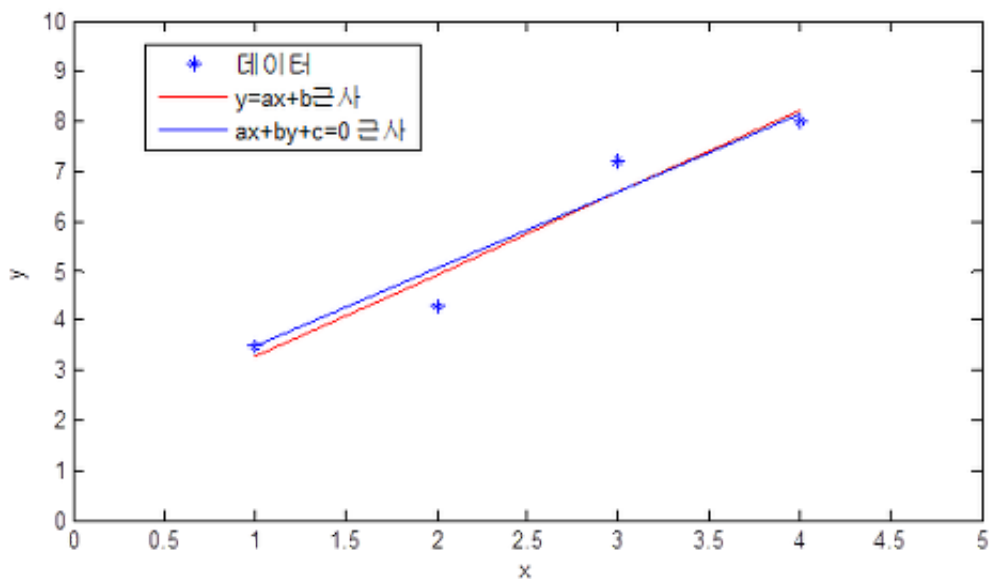
A의 특이값분해(SVD)를 해 보면 특이값이  $\sigma_1 = 13.4052$ ,  $\sigma_2 = 0.8654$ ,  $\sigma_3 = 0.3620$  로 나오고 이중 최소값인  $\sigma_3$ 에 대한 right singular vector를  $[a, b, c]$ 로 했을 때의 해는 다음과 같다.

$$0.5822x - 0.3745y + 0.7217 = 0 \quad \text{--- (17)}$$

이를 y에 대해 정리해 보면,

$$y = 1.5545x + 1.9270 \quad \text{--- (18)}$$

i) & ii) 결과를 그림으로 비교해 보면 다음과 같다.



그림으로 봐서는 어떻게 좋은 것인지 비교가 힘들기 때문에 이를 수치적으로 비교해 보자.

먼저 근사오차를  $r = y - (ax + b)$ 로 잡고 제곱합을 구해보면

i)  $y = ax + b$  모델  $\Rightarrow \sum r^2 = 0.8820$  (빨간색 그래프)

ii)  $ax + by + c = 0$  모델  $\Rightarrow \sum r^2 = 0.9345$  (파란색 그래프)

근사오차를 직선과의 수직거리로 잡은 경우에는 ( $r = |ax + by + c| / \sqrt{a^2 + b^2}$ )

i)  $y = ax + b$  모델  $\Rightarrow \sum r^2 = 0.1376$  (빨간색 그래프)

ii)  $ax + by + c = 0$  모델  $\Rightarrow \sum r^2 = 0.1311$  (파란색 그래프)

즉, y축 방향으로의 근사오차를 고려하면 i)  $y = ax + b$  모델이 좀더 낮은 에러를 보이고, 직선과의 수직거리를 근사오차로 생각하면 ii)  $ax + by + c = 0$  모델이 낮은 에러를 나타낸다.

☞ 사실 실제 계산 결과를 보기 전까지는  $y = ax + b$ 로 모델을 잡고  $AX = B$ 를 푼 결과와,  $ax + by + c = 0$ 을 모델로 잡고  $AX = 0$ 을 푼 결과가 같을지 저도 궁금했습니다만 결과는 다르게 나왔습니다. 즉,  $\|y - ax - b\|$ 를 최소화하는 해와  $\|ax + by + c\|$ 를 최소화하는 해가 다르다는 의미일 것입니다. 생각해 보면  $\|y - ax - b\|$ 는 y의 계수를 1로 고정시켜놓고 최소해를 구하기 때문에 아무런 제약이 없는  $\|ax + by + c\|$ 의 최소해보다 좋지 않은 결과를 보이는 것으로 해석할 수 있습니다.