

직교 회귀의 역학적 고찰

채 경 철*

요 약

로보트화된 좌표측정기에 사용됨에 따라, 직교 회귀가 이제는 단순히 흥미로운 수문제로 그치지 않게 되었다. 특히, 공학 전공 학생을 대상으로 하는 기초 통계학 과목에서 간단히나마 다룰만한 문제가 되었다고 본다.

직교 회귀의 핵심 개념을 간략히 정리한다. 아울러 직교 회귀의 기하학적 구조를 공학 전공 학생들이 잘 알고 있는 역학적 개념으로 설명함으로써 학생들이 쉽게 이해하고 또한 사고의 폭을 넓히는데 도움이 되고자 한다.

1. 서 론

직교 회귀에서는 수직 편차대신 직교 편차, 즉 최단거리의 최소 제곱법을 사용한다. 기초 통계학 과목에서 회귀 분석을 강의할 때 종종 학생들이 직교 회귀에 관해서 질문을 하는데 그 문제 자체가 흥미로워서 학생들의 관심을 끌기에 충분하다고 본다.

직교 회귀의 역사는 오래지만(Pearson, 1901) 근자에 와서 실제 문제에 거의 응용이 되지 않고 있었다. 고작 계량 경제학 등에서 독립 변수에 오차가 포함된 경우를 다룰 때 간혹 언급되는 정도였다. 다만 수학 또는 통계학을 전공하는 학생들만이 직교회귀의 실용성 유무를 떠나서 직접 또는(다변량 분석의 주성분 분석을 통해서) 간접으로 접해 볼 기회가 있었다고 해도 과언이 아니다. 그러다가 최근에 로보트화된(robotized) 좌표 측정기(coordinate measuring machine)의 등장과 때를 맞추어(Davison and Bland, 1986) 직교 회귀에 관한 논의가 부쩍늘고 있다(Schchat, 1985 ; Späth, 1986).

좌표 측정기는 탐침(probe)이 달린 로보트 팔이라고 할 수 있는데, 피측정 물체에 탐침이 닿으면 접촉점의 좌표가 기록된다(Schaffer, 1982). 그리고 기록된 좌표와 관련된 계산은 부착된 소형 컴퓨터가 맡는데 여기에 필요한 소프트웨어의 개발이 직교 회귀를 재등장시키는 동기가 되었다. 예를 들면,

* 한국과학기술원(대덕) 경영과학과
대전 유성구 구성동 400(305-701)

- (i) 물체 표면의 세점 이상을 측정해서 최량 적합(best fit) 평면을 계산
- (ii) 원통형 홀(hole) 또는 구형 베어링 표면상의 세점 또는 네점 이상 측정해서 최량 적합 직경을 계산
- (iii) 이 밖에도 각종 형체에 대해 최량 적합이 되는 기하적 형태의 계산등이 있다.

먼저 직교 회귀의 기본틀을 요약 정리한 다음, 대수적 해법의 기하적 구조와 해석적 해법의 역학적(mechanical) 구조를 고찰함으로써 기초 수학과 기초 역학에 익숙한 공학 전공 학생들의 이해의 폭을 넓히도록 함이 목적이다. 또한 직교 회귀의 수학적 측면을 이미 잘 알고 있는 독자도 이 기회에 눈으로 보고 느낄 수 있는 물리적 측면을 보완할 수 있겠다.

2. 직교 회귀

직교 회귀 직선 또는 평면을 구하는 문제는 다음과 같이 일반화될 수 있다. m 차원 공간에 고정된 n 개의 점으로부터의 최단거리 제곱의 합이 최소가 되는 k 차원 부분 공간(subspace)을 k 차원 직교 회귀 초평면(hyperplane)이라 하자. 여기에서 $k < m$ 이고, 또한 $k < n-1$ 이다. 편의상 n 개 점의 중심을 좌표계의 원점으로 정하고 n 개점의 좌표를 $n \times m$ 행렬 X 로 나타내기로 한다.

알려진 해법을 요약하면 다음과 같다. $X^T X$ 의 고유치(eigen value)를 큰 순서로 나열하여 $\lambda_1, \lambda_2, \dots, \lambda_m$ 이라 하고 상응하는 고유 벡터(eigen vector)를 v_1, v_2, \dots, v_m 이라 하면 v_1, v_2, \dots, v_k 가 k 차원 직교 회귀 초평면의 직교 기저(orthogonal basis)를 이룬다. 그리고 이 때 최단거리 제곱의 합은 $\lambda_{k+1} + \lambda_{k+2} + \dots + \lambda_m$ 이 된다. 예를 들어 직교 회귀 직선($k=1$)은 벡터 v_1 의 연장선이 되고, 직교 회귀 평면($k=2$)은 v_1 과 v_2 를 포함하는 평면이 된다.

해법의 유도과정은 부록 A에 상술되어 있으며 직교 회귀의 일반화에 관해서는 Kubacek(1975)와 Golub and Loan(1980)을 참조할 수 있다.

3. 기하적 구조

미식 축구공을 던질 때 보통 공이 회전하면서 나아가게 던진다. 포물선 궤적을 그리는 공의 중심을 따라가면서 공의 회전 운동을 관찰한다면 잘 던진 공은 공의 길쭉한 쪽을 회전축으로 삼아 돌지만 잘못 던진 공은 복잡한 형태의 회전 운동을 하는데 이 복잡한 형태(mode)는 세개의 기본적인 회전 형태의 중첩(superposition)으로 나타낼 수 있다. 여기에서 기본 회전 형태란 주회전축(principal axis of rotation)을 축으로 한 회전을 의미하는데 미식 축구공의 경우 잘 던진공의 회전축이 바로 첫번째 추축이고, 나머지 두 개는 대칭성에 의해 구별이 불가능하므로 첫째 축에 수직이고, 중심을 지나고, 서로 직교인 임의의 두 방향이 된다(그림 1 참조).

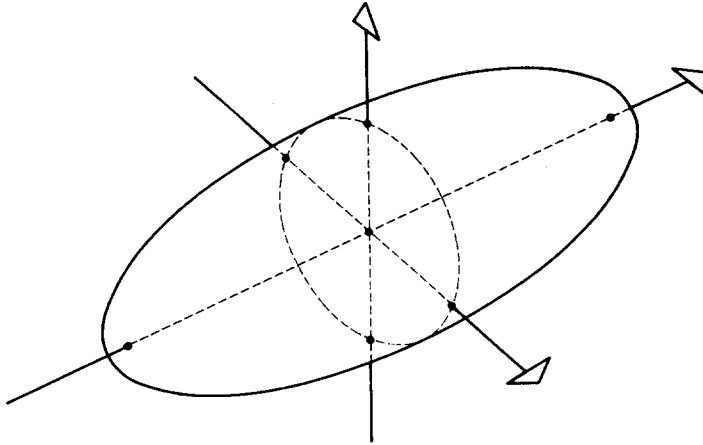


그림 1. 미식축구공의 주회전축

미식 축구공같이 넓이 또는 체적을 가진 연속체 뿐만 아니라 3차원 공간에 고정된 n 개의 점에 대해서도 세개의 주축이 정의된다. 첫째축은 n 개 점의 중심을 지나고 점들이 가장 길쭉하게 퍼진 방향이고, 둘째축은 역시 중심을 지나고 첫째축에 직교인 방향중에서 점들이 가장 길쭉하게 퍼진 방향이고, 그리고 셋째축은 중심을 지나고 첫째와 둘째축에 직교인 방향이 된다. 독자가 이미 짐작했겠지만, 주축은 다변량 관측치의 주성분(principal component)에 해당된다.

결론적으로 첫째 주축이 바로 직교회귀 직선이 되고 첫째와 둘째축이 이루는 평면이 직교회귀 평면이 되는데, 우선 주축과 직교 회귀직선의 관계를 설명하기 위해서 원점에 중심을 둔 n 개의 점을 각각 단위 질량을 가진 n 개의 질점(mass point)으로 간주하기로 한다. 그러면 n 개 질점의 관성 텐서(inertia tensor : 부록 B 참조) T 의

$$T = (\text{trace } X^T X) I - X^T X \tag{1}$$

고유벡터가 주축을 정하는데 T 가 앞에서 언급한 $X^T X$ 와 동일한 고유벡터(와 고유치)를 가지므로 첫번째 주축이 바로 직교 회귀 직선이 된다. 아울러 첫째 주축에 대한 회전 관성(rotational inertia 또는 moment of inertia : 부록 B 참조) $\lambda_2 + \lambda_3$ 가 n 개 질점으로부터 첫째 주축까지의 최단거리 제곱의 합이 된다. 다시 말해서 회전 관성을 최소화하는 회전축을 구하는 문제와 최단거리 제곱의 합을 최소화하는 직선을 구하는 문제는 같은 문제이다.

직교 회귀 평면과 주축간의 관계를 알아보기 위해 먼저 고유치의 의미를 보충하기로 한다.

위에서 첫째 주축까지의 최단거리 제곱합으로 $\lambda_2 + \lambda_3$ 를 얻었는데 여기에서 λ_2 와 λ_3 는 각각 최단거리 제곱합의 v_2 와 v_3 방향 성분이라고 볼 수 있다. 이는 피타고라스 정리에 의해 첫째 주축까지의 최단 거리 제곱이 최단거리의 v_2 방향 성분의 제곱과 v_3 방향 성분의 제곱을 합한 것과 같기 때문이다. 마찬가지로 둘째 주축까지의 최단거리 제곱 합으로 $\lambda_1 + \lambda_3$ 를 얻는데 이는 최단거리 제곱합의 v_1 방향 성분과 v_3 방향 성분의 합이 된다. 이제 v_1 과 v_2 가 이루는 평면 즉, 첫째와 둘째 주축이 이루는 평면을 고려한다. n 개 점으로부터 첫째 축까지의 최단거리 제곱합인 $\lambda_2 + \lambda_3$ 중 λ_2 의 몫은 이제 평면에 포함되어 없어지고, 역시 둘째 축에 대한 $\lambda_1 + \lambda_3$ 중 λ_1 이 없어져서 결국 λ_3 만 남게 되는데 이 값이 바로 v_1 과 v_2 가 이루는 직교 회귀 평면까지의 최단거리 제곱합임을 쉽게 알 수 있다(구체적인 사항은 부록 A 참조.)

4. 역학적 구조

3차원 공간에 고정된 n 개의 점을 벡터 $x_j(j = 1, 2, \dots, n)$ 로 나타내되 앞에서와는 달리 좌표계의 원점을 n 개 점의 중심에 두지 않는다.*

평면 S 의 방정식을 평면상의 점 x 의

$$x = a + r_1 b_1 + r_2 b_2$$

집합으로 나타내자. 여기에서 벡터 a 와 단위 벡터 b_1, b_2 는 서로 직교이고, r_1, r_2 는 스칼라 매개변수(scalar parameter)이다. 직선 방정식은 따로 정의하지 않고 위 평면 방정식에서 $r_2 = 0$ 으로 놓고 사용한다. 직교 편차를

$$f_j = x_j - (a + r_1 b_1 + r_2 b_2), \quad j = 1, 2, \dots, n$$

으로 정의한다. 여기에서 $(r_1 b_1 + r_2 b_2)$ 는 S 에의 x_j 의 투영(projection)이고 벡터 f_j 의 길이 $\|f_j\|$ 는 j 번째 고정점으로부터 S 까지의 최단거리가 된다. 따라서 n 개 점으로부터 S 에 이르는 최단거리 제곱의 합 Z 는 다음과 같이 표시된다.

$$z = \sum_{j=1}^n \|f_j\|^2 \quad (2)$$

*해법 표기상의 간편성과 유도과정의 용이성을 위해서 본문 2절과 3절 그리고 부록 A의 (A. 5)식 이하 부분에서는 n 개 점의 중심에 좌표계의 원점을 두었다. 즉, 문제에 따라 고유한 의미를 가지도록 정의되어 사용중이거나 또는 임의로 설정되어 사용중이던 좌표계를 무시하고 편의상 원점이 중심에 오도록 하는 좌표변환을 일단 시행한 다음에 문제를 다룬 셈이 되겠다. 따라서 필요에 따라서는 새 좌표계를 사용해서 얻은 결과를 원래 좌표계로 역 변환해서 문제에서 고유한 의미를 가지는 결과로 환원을 해야 되겠다. 평면 방정식을 부록 A에서는 자연방정식(intrinsic equation) 형태로 나타내었기 때문에 원점을 중심에 두는 것이 도움이 되었는데 여기서는(4절 네째줄) 평면방정식을 매개변수 방정식(parametric equation) 형태로 나타냄으로 인해 원점을 중심에 둔다고 해도 (" $a = 0$ "가 된다는 것 이외에는) 별다른 이점이 없기에 원점에 관한 제약을 더이상 두지 않는다.

Z를 a, b_1, b_2 에 대해 편미분하여 Z가 최소가 되게 하는 S에 대한 필요조건으로써 다음의 정규 방정식(normal equation)을 얻는다.

$$\sum_{j=1}^n f_j = 0 \tag{3}$$

$$\sum_{j=1}^n r_{ij} f_j = 0, i = 1, 2 \tag{4}$$

이제 직교 편차를 S에 작용하는 힘(force)이라 하면 위의 정규 방정식은 S가 평형(equilibrium) 상태에 있을 조건식이 된다. 즉, 합력(net force)이 0이라는 조건 (3)이 S의 직선 운동을 막고, 회전 효과(torque : 부록 B 참조)의 합이 0이라는 조건 (4)가 S의 회전 운동을 막는다. 아울러 S가 n개 점의 중심을 지나는 것도 조건 (3)에 기인하는 것임을 쉽게 알 수 있다.

거리에 비례하는 힘을 설명할 때 흔히 용수철을 예로 든다. 편의상, 늘이지 않은 용수철의 길이를 0이라 하고, 늘일 때 단위 길이당 한 단위의 힘을 요한다고 하자. 이와 같은 용수철의 한쪽 끝을 j번째 고정점에 매달고 용수철을 S에 직교 방향으로 늘여서 다른 쪽 끝을 S에 연결하면 이 용수철이 S를 잡아 당기는 힘은 f_j 가 될 것이다. 그리고 이 때 용수철을 늘이는데 소요된 에너지는 $1/2 \|f_j\|^2$ 이 된다. S가 움직이지 않도록 단단히 붙든 상태에서 n개의 고정점에 동일한 용수철을 하나씩 매달고 각각을 S에 직교 방향으로 늘여서 S에 연결한다. 그러면 여기에 소요된 총 에너지는 (2)의 절반 즉, Z/2가 된다. 이제 붙들고 있던 S를 움직일 수 있게 놓아주되 S에 연결된 용수철의 끝부분이 S상에서 자유로이 움직일 수 있다고 하자. 그러면 오랜 시간이 지난 뒤 S가 평형상태에 이르렀을 때 총 에너지가 최소가 된다. 다시 말해서 최단거리 제곱합이 최소가 되기 위한 조건인 정규 방정식이 에너지가 최소가 되기 위한 조건인 평형 조건에 해당된다. 참고로 덧붙이면, 에너지가 최대가 되는 상태 역시 평형 상태이긴 하지만 이 경우는 불안정(unstable) 평형 상태이므로 S를 조금만 움직여 주면 일단 비평형 상태로 되었다가 언젠가는 안정 평형 상태에 이르게 된다.

5. 예제

xy평면상의 7개 점이 아래와 같이 주어졌을 때

$$X^T = \begin{bmatrix} -5.2 & -3 & 1.2 & 3 & -1.4 & 1 & 4.4 \\ -3.6 & -4 & -3.4 & -1 & 4.8 & 3 & 4.2 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 68.8 & 38.4 \\ 38.4 & 91.2 \end{bmatrix} \text{ 또는 (1)식으로 정의된 } T = \begin{bmatrix} 91.2 & -38.4 \\ -38.4 & 68.8 \end{bmatrix} \text{ 의}$$

고유치로 $\lambda_1 = 120$ 과 $\lambda_2 = 40$ 을 얻고, 표준(normalized) 고유 벡터로 $v_1 = (0.6, 0.8)^T$ 와 $v_2 =$

$(-0.8, 0.6)^T$ 를 얻는다. 따라서 직교 회귀 직선 (또는 첫번째 주축)은 $y = (4/3)x$ 가 되고 이때 직교 편차 제곱합(또는, 단위 질량 질점일 때, 첫번째 주축에 대한 회전 관성)은 $\lambda_2 = 40$ 이 된다. 또한 힘의 평형의 관점에서 보면, 직선의 위치가 $y = (4/3)x$ 일 때

$$(f_1 \cdot \dots \cdot f_7) = (2 \ 0 \ -3 \ -3 \ 4 \ 1 \ -1)v_2$$

$$(r_{11} \cdot \dots \cdot r_{17}) = (-6 \ -5 \ -2 \ 1 \ 3 \ 3 \ 6)$$

이 되어 평형 조건 (3)과 (4)를 만족시킴을 알 수 있다(그림 2 참조).

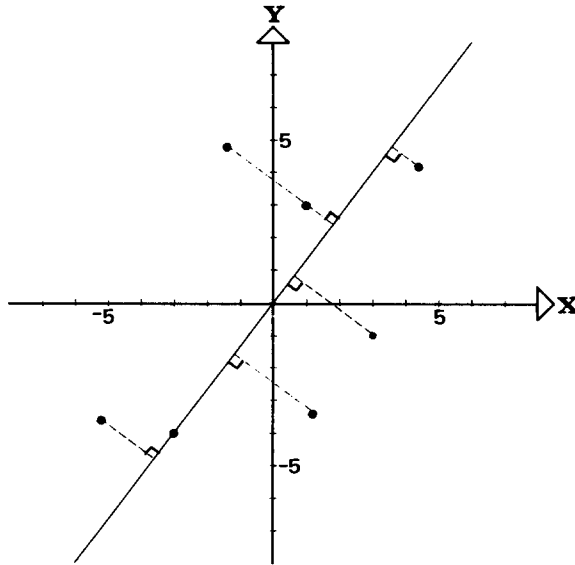


그림 2. 직교회귀직선 $y = 4/3x$

이 예제에서 사용된 수치는 아래의 방법으로 얻었다. 먼저 r_{1j} 값을 임의로 정한 다음 평형 조건 (3)과 (4)를 이용해서 (약간의 시행착오를 거쳐서) f_j 를 얻는다. 다음 (r, f) 를 원점을 축으로 시계 반대 방향으로 $\tan^{-1}(4/3)$ 라디안만큼 돌려서 위의 X^T 를 얻었는데, 임의성의 한도 내에서 7개 점의 중심이 원점에 오도록 했다. 단순 회귀 분석 강의용 예제도 이와 같은 방법으로 만들 수 있다. 예를 들어 표본 회귀 직선이 $y = 3 + 2x$ 가 되고 x 좌표의 평균치가 8이 되는 표본의 x, y 좌표로서

$$x_j = r_{1j} + 8$$

$$y_j = (3 + 2x_j) + f_j$$

로부터 $(x_1 \cdot \dots \cdot x_7) = (2 \ 3 \ 6 \ 9 \ 11 \ 11 \ 14)$ 와 $(y_1 \cdot \dots \cdot y_7) = (9 \ 9 \ 12 \ 18 \ 19 \ 26 \ 30)$ 을 얻는다. 이 예제에서의 잔차 제곱합은 앞에서 얻은 λ_2 값 즉 40이 되고, $\sum_{j=1}^7 (x_j - 8)^2 = 120$ 은 λ_1 에

해당된다. 이와 같은 단순 회귀 예제를 강의에 사용하는 이유는 회귀분석에 필요한 통계치의 대부분을 그래프로부터 바로 얻을 수 있음으로 인해 학생들이 문제의 해결과정을 직접 눈으로 보고 직관력을 키울 수 있다는 점이다.

부록 A : 직교회귀 해법의 약식 유도과정

<본문 2절 참조>

쉽게 구할 수 있는 책 중에서 직교회귀에 관해 비교적 자세히 설명되어 있는 책으로 Malinvaud (1980)을 꼽을 수 있다. 1964년에 불어판 초판이 발행되었는데 그 당시만 해도 직교회귀가 비교적 큰 비중으로 다루어지지 않았나 생각된다. 3차원* 공간에서 직교회귀 직선과 직교회귀 평면을 구하는 과정을 알기 쉽게 정리한다.

먼저 3차원 공간에 있는 임의의 직선 ($k = 1$) 또는 평면 ($k = 2$)을 아래와 같이

$$Vx = b \quad (A. 1)$$

나타내는데, 여기에서 V 는 $(3-k) \times 3$ 행렬, x 는 3×1 벡터, 그리고 b 는 $(3-k) \times 1$ 벡터이다. 평면($k = 2$)의 경우 $(3-k)$ 가 1이 되어 V 는 벡터, 그리고 b 는 스칼라가 됨으로 (A. 1)이 평면방정식이 됨을 쉽게 알 수 있다. 이 경우 V 는 평면방정식의 계수 벡터인 동시에 평면에 수직인, 따라서 평면의 방향을 나타내는 벡터가 된다. (A. 1)이 직선 ($k = 1$)인 경우에는 $(3-k)$ 개 즉 2개의 평면의 교선으로서의 직선을 정의하는 셈이 되겠다. 따라서 (A. 1)을 $(3-k)$ 개 평면방정식의 연립이라고 부를 수 있겠는데, 계산 편의를 위해서 각각의 평면방정식의 양변에 적절한 상수를 곱해서 계수를 직교정규화(ortho-normalize)한다. 다시 말해서, 일반성을 잃지 않고도 다음의 조건을 부과할 수 있는데,

$$VV^T = 1 \quad (A. 2)$$

이렇게 하면 b 는 원점으로부터 (A. 1)으로 나타내는 직선 또는 평면까지의 (최단)거리를 나타내게 된다. 즉, 평면 ($k = 2$)의 경우 스칼라 b 는 원점으로부터 이 평면까지의 거리가 되고, 직선 ($k = 1$)의 경우에는 벡터 b 의 두개 성분이 각각 원점으로부터 이 직선을 교선으로 하는 서로 직교인 두개의 평면까지의 거리가 되어 결국 원점에서 이 직선까지의 거리제곱은 벡터 b 의 두개 성분의 제곱합으로 나타낼 수 있게 된다.**

이제 3차원 공간에 고정된 n 개의 점을 벡터 $x(j = 1, 2, \dots, n)$ 로 나타내면, 위에서와 비슷한 방법으로 j 번째 점으로부터 (A. 1)으로 나타내는 직선 또는 평면까지의 거리를 $(Vx_j -$

* 가시적인 효과를 기하기 위해서 3차원 문제를 다룬다. 그렇지만 3차원 문제의 결과만 가지고도 본문 2절에 요약된 m 차원 해법을 쉽게 유추할 수 있다고 본다.

b)로 나타낼 수 있는데 이를 사용해서 극소화할 목적함수 Z 를 다음과 같이 나타낸다.

$$Z = \sum_{j=1}^n (Vx_j - b)^T (Vx_j - b) \quad (\text{A. 3})$$

다음으로, n 개 점의 중심 $\bar{x} = \sum x_j/n$ 를 사용해서 (A. 3)식을 아래와 같이 고쳐 쓴다.

$$Z = \sum_{j=1}^n (x_j - \bar{x})^T V^T V (x_j - \bar{x}) + n(V\bar{x} - b)^T (V\bar{x} - b) \quad (\text{A. 4})$$

그러면, (A. 4)식에서 b 는 둘째 항에만 있으므로 Z 가 최소가 되기 위한 조건으로 먼저 $V\bar{x} = b$ 를 얻게 되는데 이는 직교 회귀 평면 또는 직교회귀 직선이 n 개 점의 중심을 지남을 의미한다.

이 사실을 이용해서 편의상 처음부터 n 개 점의 중심을 좌표계의 원점으로 삼았을 경우 Z 는 아래와 같이 간단히 표시된다.

$$Z = \sum_{j=1}^n x_j^T V^T V x_j \quad (\text{A. 5})$$

그리고 V 의 i 번째 행을 벡터 v_i^T 로 표시하고, n 개의 점의 좌표를 $n \times 3$ 행렬 X 로 나타내어 (A. 5)와 (A. 2)를 각각 아래와 같이 고쳐 쓴다.

$$Z = \sum_{i=1}^{3-k} v_i^T X^T X v_i \quad (\text{A. 6})$$

$$v_h^T v_g = \delta_{hg} \quad (\text{A. 7})$$

이제 마지막으로 제약조건(A. 7)을 수반한 극소화 문제의 목적함수 (A. 6)에 제약조건을 포함시켜서 제약이 없는 극소화 문제의 목적함수로 바꾸고, 이를 v_i 로 미분하여 아래의 필요조건을 얻는다.**

$$(X^T X - \lambda_i I) v_i = 0, \quad i = 1, \dots, 3-k \quad (\text{A. 8})$$

여기에서 λ_i 는 Lagrange 승수(multiplier)인데 (A. 8)식이 또한 $X^T X$ 의 고유 방정식(또는 특성 방정식)이 되므로 λ_i 는 $X^T X$ 의 고유치가 된다. 아울러 v_i 는 λ_i 에 상응하는 고유벡터가 되겠다. 또한 (A. 7)과 (A. 8)로부터 아래와 같은 새로운 형태의 필요조건을 얻는데

$$v_i^T X^T X v_i = \lambda_i \quad (\text{A. 9})$$

이 식을 (A. 6)에 대입해서 필요조건을 만족시키는 목적함수로서 다음을 얻는다.

$$Z = \sum_{i=1}^{3-k} \lambda_i \quad (\text{A. 10})$$

** 일부 생략된 유도과정과 더욱 구체적인 설명은 Malinvaud를 참조하기 바람.

따라서 Z의 최소치는 $X^T X$ 의 세계의 고유치 중에서 가장 작은 것 $(3-k)$ 개를 더한 것이 된다.

다시 말해서 $X^T X$ 의 고유치를 큰 순서로 $\lambda_1, \lambda_2, \lambda_3$ 라 할 때, n 개 점으로부터 직교 회귀 평면($k=2$)까지의 거리제곱의 합은 가장 작은 $(3-k)$ 개 고유치의 합 즉 λ_3 가 되고 이 때 직교회귀 평면은 v_3 와 수직을 이루는 평면 즉 v_1 과 v_2 가 이루는 평면이 된다. 또한 직교회귀 직선($k=1$)까지의 거리제곱합은 가장 작은 $(3-k)$ 개 고유치의 합 즉 $(\lambda_2 + \lambda_3)$ 가 되고 이 때 직교회귀 직선은 v_2 에 수직인 평면과 v_3 에 수직인 평면이 이루는 교선 즉 v_1 의 연장선이 될 것이다.

<본문 3절 참조>

고유치의 의미를 더욱 분명하게 하기 위해서 j 번째 점의 v_1, v_2, v_3 세 방향 성분을 각각 r_{1j}, r_{2j}, r_{3j} 라 하자(그림 3 참조). 그러면 v_3 에 수직인 직교회귀 평면까지의 거리제곱합 λ_3 는 $\sum_{j=1}^n r_{3j}^2$ 이 되고 v_1 의 연장선인 직교회귀 직선까지의 거리제곱합($\lambda_2 + \lambda_3$)는 $\sum (r_{2j}^2 + r_{3j}^2)$ 이 된다.

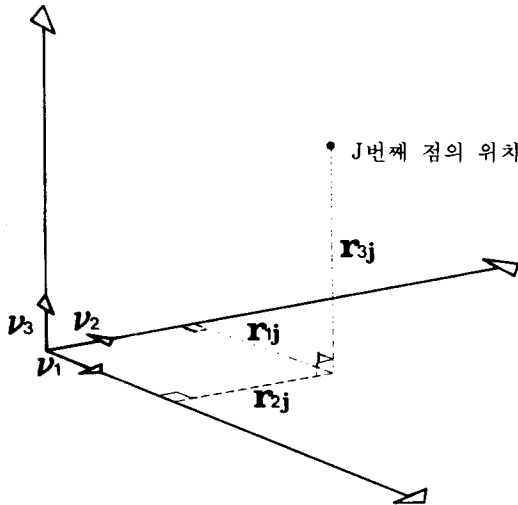


그림 3. 직교편차와 주축방향성분

한편, $k=0$ 인 특수한 경우에는 다음아닌 n 개 점의 중심이 바로 n 개 점으로부터의 (최단) 거리 제곱의 합이 최소가 되는 “점 ($k=0$)”이 되는데, 이 때 거리제곱합($\lambda_1 + \lambda_2 + \lambda_3$)는 $\sum (r_{1j}^2 + r_{2j}^2 + r_{3j}^2)$ 이 될 것이다. 이상을 종합하여 다음을 얻는다.

$$\lambda_i = \sum_{j=1}^n r_{ij}^2, \quad i = 1, 2, 3$$

부록 B : 역학 용어 해설

회전효과(본문 4절), 회전관성(3절), 관성텐서(3절)의 순서로 해설한다.

단단한 물체(rigid body)의 운동은 직선운동과 회전운동의 중복(superposition)으로 나타낸다. 본문에서 예로든 미식축구공이 잘 던져졌을 경우, 이 공의 운동은 공의 중심의 직선운동과 중심을 지나는 첫째 주회전축을 축으로 한 회전운동의 중복이 된다. 이 때 물론 공의 중심이 그리는 포물선 궤적은 중심의 수평 등속운동과 수직 가속운동 즉 수평, 수직 두 방향의 직선운동의 중복이다. 아울러 잘못 던진 미식축구공의 회전운동 역시 세 방향의 주축을 축으로 한 회전운동의 중복이다.

먼저 직선운동을 별도로 다루기 위해서 편의상 한개의 질점을 고려한다. 뉴턴의 제2법칙에 따르면 질점은 질점에 가해진 힘에 비례하는 가속운동을 하는데 이 때의 비례상수를 우리는 관성질량이라고 부른다. 따라서 같은 크기의 힘(f)이 작용하더라도 질점의 가속도(a)는 관성질량(M)에 반비례하게 되는데

$$f = Ma \tag{B. 1}$$

이로부터 M 을 “관성”질량이라 부르는 이유를 찾을 수 있을 것이다.

다음, 회전운동을 별도로 다루기 위해서 시소(seesaw)틀을 고려한다. 시소틀에 수직방향으로 힘을 가하면 그 힘에 비례하는 회전 가속운동을 한다. 그런데 같은 크기의 힘이라도 작용점이 회전축에서 멀리 떨어져 있을수록 회전효과가 크다는 것을 우리는 잘 알고 있다. 따라서 **회전효과(torque)**는 작용하는 힘(의 시소틀에 대한 수직성분) f 와 작용점의 축으로부터의 거리 r 의 곱 즉, rf 로 정의한다(본문 (4)식 참조)

이제 회전운동에 대해서 직선운동의 (B. 1)에 대응하는 관계를 알아보기로 한다. (B. 1)의 좌변의 f 대신 회전효과 rf 를 대응시키고, 우변의 선(線) 가속도 a 를 각(角) 가속도 $\alpha (=a/r)$ 로 대체하면, 회전효과와 각가속도간의 비례상수로서 (B. 1)의 관성질량에 해당되는 회전관성으로서 Mr^2 를 얻는다.

시소틀 문제를 떠나서, 일반적으로 공간에 고정된 n 개 질점의 질량이 M_1, M_2, \dots, M_n 이고 특정축으로부터의 (최단)거리가 r_1, r_2, \dots, r_n 이라면 이 시스템의 **회전관성**은 $\sum_{j=1}^n M_j r_j^2$ 이 된다.

(유의사항 : 본문에서는 편의상 모든 질량을 단위질량으로 간주했음)

마지막으로 **관성텐서**란 한 마디로 회전관성을 현재 사용중인 특정 좌표계로 나타낸 것이라고 볼 수 있다(본문 (1)식 참조). 원래 관성텐서는 각운동량과 각속도간의 관계*를 나타내기 위해서 정의된 약간 추상적인 개념이긴 하지만 최소한 대각(diagonal)성분은 각각 x 축, y 축, z 축에 대한 회전관성이라는 의미가 있다. 반면에 관성의 곱(product of inertia)이라 불리는 비대각 성분은 각각 두개의 축에, 예를 들어 x 축과 y 축에, 복합적으로 관련된 성분이라는 정도 이상의 의미를 부여하기는 어렵고 다만 통계학에서의 공분산(covariance)에 해

당된다고 할 수 있는 정도이다. 좌표변환을 한다고 해서 기하적 대상(점, 선, 벡터 등)이 움직이는 것은 아니고 다만 고정된 기하적 대상의 위치를 수치로 표시하는 방법(즉 좌표계)만 바꾸는 것이라는 점을 염두에 두고, 관성텐서의 세개의 고유벡터를 새로운 좌표축으로 삼는 좌표변환을 하는 경우에 새 좌표로 표시된 관성텐서의 비대각성분은 모두 0이 되고 대각성분은 각 고유벡터 즉 주회전축에 대한 회전관성이 된다는 점을 덧붙인다.

* 선운동량 벡터는 선속도 벡터와 같은 방향이고 이 때의 비례상수는 스칼라인 관성질량이다. 반면에, 각운동량 벡터는 일반적으로 각속도 벡터와 같은 방향이 아님으로 각운동량을 비례상수의 행렬이라고 할 수 있는 관성텐서와 각속도의 곱으로 나타낸다.

참 고 문 헌

- (1) Davison, R. and Bland, J. A. (1986). "Generalized Regression for CMMs", International Journal of Mathematical Education in Science and Technology, 17, 305~309.
- (2) Golub, G. H. and Loan, C. F. van (1980). "An Analysis of the Total Least Squares Problem", SIAM Journal on Numerical Analysis, 17, 883~893.
- (3) Kubacek, L. (1975). "On a Generalization of the Orthogonal Regression", Aplikace Matematiky, 20, 87~95.
- (4) Malinvaud, E. (1980). Statistical Methods of Econometrics, North-Holland.
- (5) Pearson, K. (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space", Philosophical Magazine, ser. 6, 2, 559~572.
- (6) Schaffer, G. (1982). "Taking the Measure of CMMs", American Machinist, October, 145~160.
- (7) Schuchat, A. (1985). "Generalized Least Squares and Eigenvalues", The American Mathematical Monthly, 92, 656~659.
- (7) Späth, H. (1986). "Orthogonal Least Squares Fitting with Linear Manifolds", Numerische Mathematik, 48, 441~445.

Mechanics of Orthogonal Regression

Kyung C. Chae*

Abstract

Orthogonal regression is reviving due to the current development of software to control robotized coordinate measuring machines. Known features of orthogonal regression are summarized and geometry of the problem is interpreted in terms of equilibrium of force. A class example is given.

*Department of Management Science,
Korea Institute of Technology
400, Kusong-dong, Yusong-gu
Taejon, 305-701