[2008][10-1]

Computer aided ship design

Part 3. Optimization Methods

November 2008

Prof. Kyu-Yeul Lee

Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

Advanced
Ship
Design
Automation
Laboratory

Ch4. 쿤-터커(Kuhn-Tucker) 정리를 이용한 최적성 조건

4.1 최적성 조건을 이용한 최적해의 도출

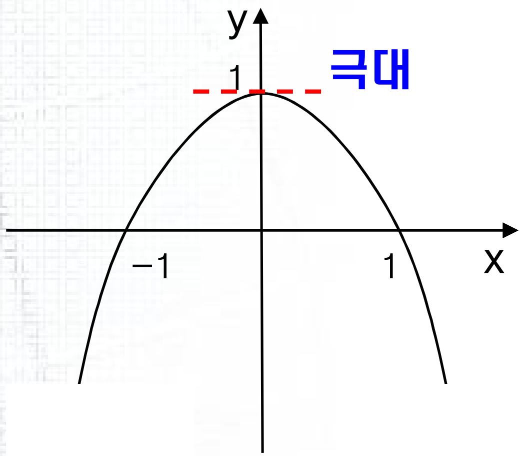
4.2 라그란지 승수(Lagrange multiplier)를 도입한 최적화 기법

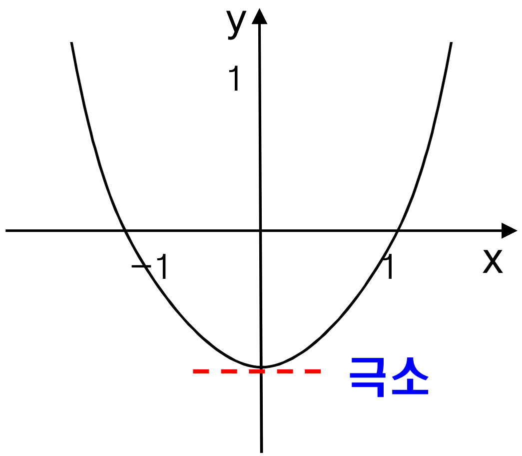
Advanced
Ship
Design
Automation
Laboratory

4.1 최적성 조건을 이용한 최적해의 도출

1변수 함수의 최적성 조건
- 함수의 극대, 극소(고교 과정 Review)

- 고등학교 수학의 정석(수학 II) Review
  - 수학의 정석 "6. 극대, 극소와 미분" (p. 104)





1) 극대값 : $x = x^*$에서 연속 함수 $f(x)$가 **증가에서 감소** 상태로 변함
2) 극소값 : $x = x^*$에서 연속 함수 $f(x)$가 **감소에서 증가** 상태로 변함

$$f'(x^*) = 0$$

($x = x^*$에서 극대 또는 극소값을 가질 필요 조건)

1변수 함수의 최적성 조건
- 1계 필요 조건(1)

- 변수가 하나일 때, 극값을 가질 필요 조건 : $f'(x^*) = 0$

pf) 주어진 점 x^*에서 $f(x)$의 테일러 급수는 다음과 같다.

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 + R$$

나머지항(Remainder)
: x가 x^*에 충분히
가까우면 그 값이 매우 작음

$x - x^* = d$ 라고 놓으면

$$f(x) = f(x^*) + f'(x^*)d + \frac{1}{2}f''(x^*)d^2 + R$$

함수 값의 변화량 $f(x) - f(x^*) = \Delta f(x)$

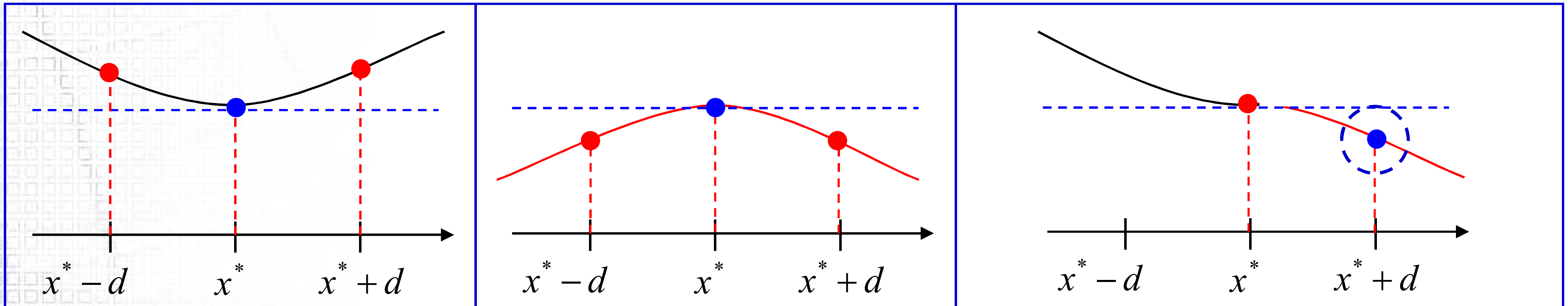
$$\Delta f(x) = f'(x^*)d + \frac{1}{2}f''(x^*)d^2 + R$$

1변수 함수의 최적성 조건
- 1계 필요 조건(2)

$$\Delta f(x) = f(x) - f(x^*) = f'(x^*)d + \frac{1}{2} f''(x^*)d^2 + R$$ ➡

의미 : x^*에서의 함수 값이 $x^* \pm d$에서의 함수 값보다 작으면 x^*는 국부적 후보 최소점이 될 가능성이 있음

$x = x^*$ 에서 국부적 후보 최소점이기 위해서는 $\Delta f \geq 0$



$f(x^*)$ 에서 최소값

$f(x^*)$ 에서 최대값

$f(x^*)$ 보다 함수 값이 작은 점(x^*+d)이 존재하므로 x^*는 최소점이 아님

따라서, $d(= x - x^*)$ 의 부호에 상관없이 $\Delta f \geq 0$를 만족하려면 $f'(x^*) = 0$ 이어야 함

이와 유사한 개념으로 $x = x^*$ 에서 국부적 후보 최대점이 되기 위해서는 $\Delta f \leq 0$이어야 하며, $d(= x - x^*)$의 부호에 상관없이 $\Delta f \leq 0$를 만족하려면 $f'(x^*) = 0$이어야 함

cf) $f'(x^*) = 0$를 만족하는 점 : 국부적으로 최소, 최대, 또는 변곡점
총칭하여 상점(Stationary point)이라고 함

1변수 함수의 최적성 조건
- 충분 조건과 2계 필요 조건

$$\Delta f(x) = f(x) - f(x^*) = f'(x^*)d + \frac{1}{2} f''(x^*)d^2 + R$$

$x = x^*$ 에서 국부적 후보 최소점이기 위해서는 $\Delta f \geq 0$

따라서, $d(= x - x^*)$ 의 부호에 상관없이 $\Delta f \geq 0$를 만족하려면 $f'(x^*) = 0$ 이어야 함

- 상점($f'(x^*) = 0$를 만족하는 점) 중 어느 점이 최소점인지 결정하는 방법

상점이므로, $f'(x^*) = 0$. 따라서, $\Delta f(x) = \frac{1}{2} f''(x^*)d^2 + R$

2차항이 다른 모든 고차항에 비해 지배적인 항이므로,

$d \neq 0$에 대하여 다음 조건을 만족하면 항상 $\Delta f \geq 0$

$f''(x^*) > 0$ (충분 조건(Sufficient condition))

cf) $f''(x^*) = 0$ 인 경우 국부적으로 최소점인지 알 수 없다.

∵ 1계, 2계 미분이 모두 0이므로, 국부적으로 최소이기 위해서는
3계 미분계수는 0이 되어야 하고, 4계 미분계수가 양수(>0)여야 한다.
즉, R의 부호에 따라 달라지게 된다.

- 2계 필요 조건 :

국부적으로 최소값을 가지면, $f''(x^*) \geq 0$

[※ 단, 역은 성립하지 않는다.]

1변수 함수의 최적성 조건(요약)

- 충분 조건, 필요 조건, 필요 충분 조건

(고등학교 수학의 정석 '명제와 조건' 참고)

1) A는 B의 충분 조건이다.

"A이면 B이다." ($A \Rightarrow B$)

2) C는 D의 필요 조건이다.

"D이면 C이다." ($C \Leftarrow D$)

3) E는 F의 필요 충분 조건이다.

"E이면 F이고, F이면 E이다." ($E \Leftrightarrow F$)

- 1계 필요 조건(1st Order necessary condition)

x^*가 국부적 후보 최소점이면, $f'(x^*) = 0$

cf) $f'(x^*) = 0$ 이면, x^*가 상점(극소, 극대, 변곡점) 이다.

- 충분 조건(Sufficient condition)

상점($f'(x^*) = 0$)일 경우

$f''(x^*) > 0$ 이면 x^*가 국부적 최소점이다.

- 2계 필요 조건(2nd Order necessary condition)

x^*가 국부적 최소점이면, $f''(x^*) \geq 0$

2변수 함수의 테일러 전개(Taylor Series Expansion)(1)

2변수 함수 $f(x_1,x_2)$에 대한 점 (x_1^*,x_2^*)에서의 테일러 전개식

$$f(x_1,x_2)=f(x_1^*,x_2^*)+\frac{\partial f}{\partial x_1}(x_1-x_1^*)+\frac{\partial f}{\partial x_2}(x_2-x_2^*)$$

$$+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}(x_1-x_1^*)^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}(x_1-x_1^*)(x_2-x_2^*)+\frac{\partial^2 f}{\partial x_2^2}(x_2-x_2^*)^2\right)+R$$

⬇ 각 항을 다시 표현하면

$$\frac{\partial f}{\partial x_1}(x_1-x_1^*)+\frac{\partial f}{\partial x_2}(x_2-x_2^*)=\begin{bmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\end{bmatrix}^T\begin{bmatrix}x_1-x_1^*\\ x_2-x_2^*\end{bmatrix}=\underline{\nabla f(\mathbf{x}^*)^T(\mathbf{x}-\mathbf{x}^*)}$$

$$\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}(x_1-x_1^*)^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}(x_1-x_1^*)(x_2-x_2^*)+\frac{\partial^2 f}{\partial x_2^2}(x_2-x_2^*)^2\right)=\frac{1}{2}\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2}(x_1-x_1^*)+\frac{\partial^2 f}{\partial x_2\partial x_1}(x_2-x_2^*) & \frac{\partial^2 f}{\partial x_1\partial x_2}(x_1-x_1^*)+\frac{\partial^2 f}{\partial x_2^2}(x_2-x_2^*)\end{bmatrix}\begin{bmatrix}x_1-x_1^*\\ x_2-x_2^*\end{bmatrix}$$

$$=\frac{1}{2}\begin{bmatrix}x_1-x_1^* & x_2-x_2^*\end{bmatrix}\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2}\end{bmatrix}\begin{bmatrix}x_1-x_1^*\\ x_2-x_2^*\end{bmatrix}$$

$$=\underline{\frac{1}{2}(\mathbf{x}-\mathbf{x}^*)^T\mathbf{H}(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*)}$$

$$f(\mathbf{x})=f(\mathbf{x}^*)+\underline{\nabla f(\mathbf{x}^*)^T(\mathbf{x}-\mathbf{x}^*)}+\underline{\frac{1}{2}(\mathbf{x}-\mathbf{x}^*)^T\mathbf{H}(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*)}+R$$

2x2 Matrix의 원소

$$\left(\ \mathbf{x}=(x_1,x_2)^T\ ,\ \mathbf{x}^*=(x_1^*,x_2^*)^T\ ,\ \boxed{\mathbf{H}\in M_{2\times 2}}\ \right)$$

2변수 함수의 테일러 전개(2)

- 2변수 테일러 전개식의 행렬 표기법

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T (\mathbf{x}-\mathbf{x}^*) + \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^*\right)^T \mathbf{H}(\mathbf{x}^*)\left(\mathbf{x}-\mathbf{x}^*\right) + R$$

2x2 Matrix의 원소

$$\left(\mathbf{x} = (x_1, x_2)^T,\ \mathbf{x}^* = (x_1^*, x_2^*)^T,\ \boxed{\mathbf{H} \in M_{2\times 2}} \right)$$

- 다변수 함수의 테일러 전개식의 행렬 표기법 : 2변수 전개식과 동일

$\mathbf{x},\ \mathbf{x}^*, \nabla f$: n 차원 Vector

$\mathbf{H} \in M_{n\times n}$

- $\mathbf{x}-\mathbf{x}^* = \mathbf{d}$ 로 정의 하면, 다음과 같이 표현 가능

$$f(\mathbf{x}^* + \mathbf{d}) = f(\mathbf{x}^*) + \boxed{\nabla f(\mathbf{x}^*)^T}\,\mathbf{d} + \boxed{\frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d}} + R$$

$\boxed{\nabla f(\mathbf{x}^*)^T = 0,\ \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0}$ ➡ $\mathbf{x} = \mathbf{x}^*$에서 최소이기 위한 충분 조건

[참고] 헷세 행렬(Hessian Matrix)

- 헷세 행렬(Hessian Matrix) : Gradient Vector를 한번 더 미분하면(각각, x_i에 대하여 미분), 함수 $f(x)$에 대한 2계 편도함수의 행렬을 얻게 되는데, 이를 헷세 행렬이라고 정의함.
  즉, Gradient Vector의 각 성분을 $x_1, x_2, \cdots, x_n$ 으로 미분하면 다음을 얻는다.

$$\frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$$\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^T$$

: n개의 변수를 가지는 column Vector(열벡터)

- 헷세 행렬은 보통 H 또는 $\nabla^2 f$ 로 표시

$$\mathbf{H} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right] \quad (i = 1,2,\cdots,n \; ; \; j = 1,2,\cdots,n)$$

- 헷세 행렬의 대칭성

$$\frac{\partial f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j \partial x_i}$$

따라서, 헷세 행렬은 항상 **대칭 행렬**

2차 형식(Quadratic Form)

- 2차 형식 : 한 변수의 제곱항과 서로 다른 두 변수의 곱항(다변수 곱항)의 합으로 이루어진 함수

ex) $F(x_1, x_2, x_3) = \frac{1}{2}\left(2x_1^2 + 2x_1x_2 + 4x_1x_3 - 6x_2^2 - 4x_2x_3 + 5x_3^2\right)$

2차 형식은 다음과 같이 행렬로 표현이 가능하다.

A : 대칭 행렬
(Symmetric Matrix)

$$F\left(x_1, x_2, x_3\right) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 2 \\ 1 & -6 & -2 \\ 2 & -2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2}\mathbf{x}^T\boxed{\mathbf{A}}\mathbf{x} \Leftrightarrow \frac{1}{2}\mathbf{d}^T\mathbf{H}\mathbf{d}$$

- 대칭행렬 A의 구성 방법 (A의 (i,j) 성분을 a_{ij}로 나타냄)
  1) 대각 성분은 제곱항의 계수 $a_{ii} = \left(x_i^2 \text{의 계수}\right)$
  2) 대각 성분 이외의 성분은 다변수 곱항 계수의 ½ $a_{ij} = \left(x_i x_j \text{의 계수}\right) \times \frac{1}{2}$

2차 형식의 형태

- 2차 형식의 정의

1) 양정 형식(Positive Definite)

: $\mathbf{x} \neq 0$ 인 모든 x에 대하여 $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$

2) 양반정 형식(Positive Semidefinite)

: 모든 x에 대하여 $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$이고

$\mathbf{x}^T\mathbf{A}\mathbf{x} = 0$ 인 $\mathbf{x} \neq 0$ 가 존재

3) 음정 형식(Negative Definite)

: $\mathbf{x} \neq 0$ 인 모든 x에 대하여 $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$

4) 음반정 형식(Negative Semidefinite)

: 모든 x에 대하여 $\mathbf{x}^T\mathbf{A}\mathbf{x} \leq 0$

5) 부정 형식(Indefinite)

: 어떤 x에 대해서는 양수 또는 음수

- 2차 형식 정의의 이용

① 1변수 함수의 최소 조건

$f'(x^*) = 0$ 인 x^*(상점)에서

$f''(x^*) > 0$ 이면

x^*는 국부적 최소점이다.

② 다변수 함수의 최소 조건

$\nabla f(\mathbf{x}^*)^T = 0 \Rightarrow \nabla f(\mathbf{x}^*) = 0$ 인

$\mathbf{x}^*$(상점)에서 $\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$이면

즉, $\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$ 가 **양정 형식**이면

$\mathbf{x}^*$는 국부적 최소점이다.

여기서, $\mathbf{H}(\mathbf{x}^*)$가 양정 행렬이면

$\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$

참고: KREYSZIG E., Advanced Engneering Mathematics, WILEY, 2006, 8.4. Eigenbasis. Diagonalization. Quadratic forms.

2차 형식 행렬의 형태 결정을 위한 고유치 검사

2차 형식 $F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{A}\mathbf{x}$ 에 관련된 대칭 $n \times n$ 행렬 $\mathbf{A}$의 n개의 고유치를 $\lambda_i, i = 1,\ldots,n$ 이라고 가정

1) $F(\mathbf{x})$가 양정(Positive Definite)이기 위한 필요 충분 조건: 모든 고유치가 양수

$$\lambda_i > 0, i = 1,\ldots,n$$

2) $F(\mathbf{x})$가 양반정(Positive Semidefinite)이기 위한 필요 충분 조건: 모든 고유치가 음수가 아님

$$\lambda_i \geq 0, i = 1,\ldots,n$$

3) $F(\mathbf{x})$가 음정(Negative Definite)이기 위한 필요 충분 조건: 모든 고유치가 음수

$$\lambda_i < 0, i = 1,\ldots,n$$

4) $F(\mathbf{x})$가 음반정(Negative Semidefinite)이기 위한 필요 충분 조건: 모든 고유치가 양수가 아님

$$\lambda_i \leq 0, i = 1,\ldots,n$$

5) 적어도 하나의 λ_i가 음수이고 또 적어도 하나의 λ_i가 양수이면 $F(\mathbf{x})$ 는 부정(Indefinite)임

고유치 검사에 의한 2차 형식의 행렬의 형태 결정 예제

<u>고유치 구하는 방법</u>

$$\mathbf{Av} = \lambda\mathbf{v} \quad \Rightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = 0 \quad \Rightarrow \quad \det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

다음 행렬의 고유치를 구하여 행렬의 형태를 결정하시오.

$$\mathbf{A} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

$$\det\begin{bmatrix} 4-\lambda & 2 & 2 \\ 2 & 4-\lambda & 2 \\ 2 & 2 & 4-\lambda \end{bmatrix} = (2-\lambda)^2(8-\lambda) = 0$$

$$\therefore \lambda = 2(\text{중근}),\, 8$$

모든 고유치가 양수이므로 주어진 행렬은 양정(Positive Definite) **행렬임**

다변수 함수의 후보 최적성 조건(요약)

- n개의 변수로 이루어진 다변수 함수 $f(\mathbf{x})$ 에 대한 테일러 전개식

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} + R$$

- 함수의 변화량으로 위 식을 다시 쓰면,

$$\Delta f = \nabla f(\mathbf{x}^*)^T \mathbf{d} + \frac{1}{2}\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} + R$$

- $\mathbf{x}^*$가 국부적으로 후보 최소점이면, $\Delta f \geq 0$이어야 함

1) x*가 국부적 후보 최소점일 필요 조건 :

➡ $\nabla f(\mathbf{x}^*)^T = 0$ 즉, $\dfrac{\partial f(\mathbf{x}^*)}{\partial x_i} = 0$, $(i = 1,2,\cdots n)$

2) $\nabla f(\mathbf{x}^*)^T = 0 \Rightarrow \nabla f(\mathbf{x}^*) = 0$인 상점에서 x*가 국부적 최소점일 조건 :

➡ $\mathbf{d}^T \mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$ ➡ 위 조건은 $\mathbf{H}(\mathbf{x}^*)$ 가 양정 행렬일 경우 성립한다.

후보 최적성 필요 조건을 이용한 등호 제약 최적화 문제의 해법

Optimum Problem

$$\textit{Minimize} \quad f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$

$$\textit{Subject to} \quad h(x_1, x_2) = x_1 + x_2 - 2 = 0$$

Solution

x_2를 x_1의 양함수(explicit function)로 표현하면,

$$x_2 = \Phi(x_1) = -x_1 + 2$$

$$f(x_1, \Phi(x_1)) = (x_1 - 1.5)^2 + (-x_1 + 2 - 1.5)^2$$

$$\frac{df}{dx_1} = 2(x_1 - 1.5) - 2(-x_1 + 0.5) = 0$$

$$\Rightarrow x_1 = 1$$

$$\Rightarrow x_2 = -x_1 + 2 = 1$$

$$\frac{d^2 f}{dx_1^{\,2}} = 4 > 0$$

$\therefore (x_1,\ x_2) = (1,\ 1)$: 국부적 최소점

4.2 라그란지 승수(Lagrange multiplier)를 도입한 최적화 기법

Advanced
Ship
Design
Automation
Laboratory

최적화 문제의 분류와 해법

	비제약 최적화 문제		제약 최적화 문제		
	선형	비선형	선형	비선형	
목적 함수 (예시)	minimize $f(\mathbf{x})$ $f(\mathbf{x}) = x_1 + 2x_2$	minimize $f(\mathbf{x})$ $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$	minimize $f(\mathbf{x})$ $f(\mathbf{x}) = x_1 + 2x_2$	minimize $f(\mathbf{x})$ $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$	minimize $f(\mathbf{x})$ $f(\mathbf{x}) = x_1^2 + x_2^2 - 3x_1x_2$
제약 조건 (예시)	없음	없음	$h(\mathbf{x}) = x_1 + 5x_2 = 0$ $g(\mathbf{x}) = -x_1 \le 0$	$h(\mathbf{x}) = x_1 + 5x_2 = 0$ $g(\mathbf{x}) = -x_1 \le 0$	$g_1(\mathbf{x}) = \frac{1}{6}x_1^2 + \frac{1}{6}x_2^2 - 1.0 \le 0$ $g_2(\mathbf{x}) = -x_1 \le 0$
국부 최적화 방법	① 직접 탐사법 - Hooke&Jeeves - Nelder&Mead ② Gradient 방법 - Steepest Descent방법[4] - 공액 경사도 방법[4](Conjugate Gradient 방법) - Newton 방법[5] - Davidon-Fletcher-Powell(DFP) 방법[5] - Broyden-Fletcher-Goldfarb-Shanno(BFGS) 방법[5]		Penalty Function으로 해를 구할 수 있으나 일반적으로 선형 계획법을 사용	- Penalty Function[1]을 구성한 후 비제약 최적화 문제로 변환한 후 해를 구함	
			- Simplex 방법 (선형 계획법, Linear Programming)	- 2차 계획 법 (Quadratic Programming)	- 근사화 방법 - SLP 선형 계획 문제[2]로 근사화 후 개선된 탐색점을 찾고, 그 점에서 다시 선형 계획 문제를 푸는 방법
					- 근사화 방법 - SQP 2차 계획 문제[3]로 근사화 후 개선된 탐색점을 찾고, 그 점에서 다시 2차 계획 문제를 푸는 방법
전역 최적화 방법	Genetic Algorithms(GA), Simulated Annealing, etc.				

1) Penalty Function
제약 조건의 위배량을
원래 목적 함수에 더한
수정된 목적 함수

2) 선형 계획 문제
(Linear Programming Problem)
목적함수: 1차 형식
제약조건: 1차 형식

3) 2차 계획 문제
(Quadratic Programming Problem)
목적함수: 2차 형식
제약조건: 1차 형식

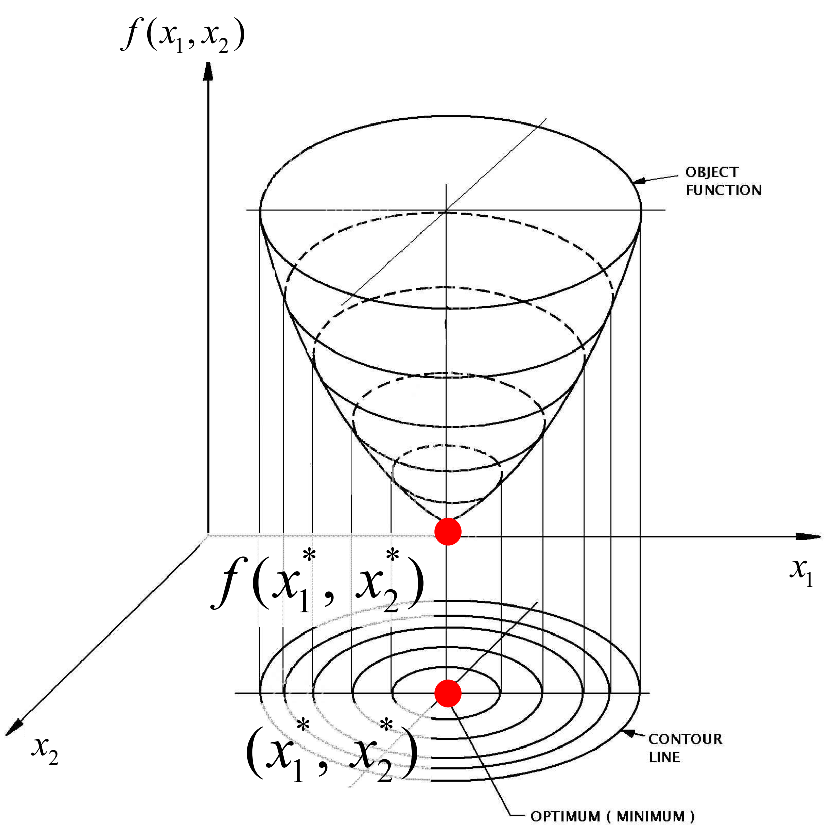
4) Gradient 방법 중
함수의 1차 미분만을
고려하는 방법

5) Gradient 방법 중
함수의 2차 미분까지
고려하는 방법

함수의 상점(Stationary Point)을 찾는 방법

Given: minimize $f(x_1, x_2)$

Find: **상점** (x_1^*, x_2^*)



- 어떤 점 (x_1^*, x_2^*)에서 미소변위(dx_1, dx_2) 만큼 이동한 점에서의 함수값의 변화량 df는 다음과 같다.

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2$$

모든 방향으로의 함수값의 변화량 df가 0인 점을 **상점(Stationary Point)**이라고 하며, 최소점, 최대점, 안장점(Saddle Point)은 상점에 속한다.

참고: 일반적인 공학적 최적화 문제에서는 목적 함수의 최적값(Optimum Value)보다는 최적점(Optimum Point)이 더 중요하다.
[예시] 건조비를 최소화 하는 선박의 주요치수(L, B, D, C_B)가 건조비 자체보다 중요한 개념이다.

제약조건이 없는 경우의 함수와 상점(Stationary Point)

Given: $minimize\ f(x_1, x_2, x_3)$

Find: **상점** (x_1^*, x_2^*, x_3^*)

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3$$

상점에서의 df **는 0이다.**
dx_1, dx_2, dx_3**는 임의의 값이라고 하면** df **가 항상 0**
이 되기 위해서는

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial x_3} = 0$$

$$\Rightarrow \nabla f = 0$$

이 되어야 한다.

제약조건이 있는 경우의 함수와 상점(Stationary Point)

Given: $minimize\ f(x_1, x_2, x_3)$

$$h(x_1, x_2, x_3) = 0$$

Find: **상점** (x_1^*, x_2^*, x_3^*)

1. 함수 h (제약조건)를 x_1에 대해서 정리
2. 함수 f에 x_1 을 대입하여 정리하고 f의 전미분이 0임을 이용하여 상점 구하기

제약 조건 h 를 하나의 변수로서 정리하기 어려울 수 있으며, 정리할 변수를 선택하는 것도 쉽지 않을 수 있다.

하나의 변수로 정리하지 않고, 상점을 구하는 방법은 없을까?

예시) 제약 조건이 다음과 같은 경우 어느 한 변수로 정리하기 어려움

$$h(x_1, x_2, x_3) = \tan x_1 + \cos x_2 + e^{x_3} = 0$$

상점에서의 df 는 0이다.

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3 = 0 \quad \text{-----} ①$$

$h(x_1,x_2,x_3)=0$ 이므로 dh 는 0이다.

$$dh = \frac{\partial h}{\partial x_1} dx_1 + \frac{\partial h}{\partial x_2} dx_2 + \frac{\partial h}{\partial x_3} dx_3 = 0 \quad \text{-----} ②$$

식 ①, ②는 모두 0이므로 다음 식이 항상 성립한다.

$$df + \lambda \cdot dh = 0$$

제약조건이 있는 경우의 함수와 상점(Stationary Point)

Given: $minimize\ f(x_1, x_2, x_3)$

$h(x_1, x_2, x_3) = 0$

Find: **상점** (x_1^*, x_2^*, x_3^*)

① $df = \dfrac{\partial f}{\partial x_1}dx_1 + \dfrac{\partial f}{\partial x_2}dx_2 + \dfrac{\partial f}{\partial x_3}dx_3 = 0$

② $dh = \dfrac{\partial h}{\partial x_1}dx_1 + \dfrac{\partial h}{\partial x_2}dx_2 + \dfrac{\partial h}{\partial x_3}dx_3 = 0$

- **제약조건 h가 있기 때문에 dx_1, dx_2, dx_3가 1차 독립이 아님.**

$$df + \lambda \cdot dh = 0$$

위 식을 정리 하면

$$\frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 + \lambda\left(\frac{\partial h}{\partial x_1}dx_1 + \frac{\partial h}{\partial x_2}dx_2 + \frac{\partial h}{\partial x_3}dx_3\right) = 0$$

$$\left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right)dx_1 + \left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right)dx_2 + \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right)dx_3 = 0$$

제약조건이 있는 경우의 함수와 상점(Stationary Point)

Given: $minimize\ f(x_1, x_2, x_3)$

$h(x_1, x_2, x_3) = 0$

Find: **상점** (x_1^*, x_2^*, x_3^*)

① $df = \dfrac{\partial f}{\partial x_1}dx_1 + \dfrac{\partial f}{\partial x_2}dx_2 + \dfrac{\partial f}{\partial x_3}dx_3 = 0$

② $dh = \dfrac{\partial h}{\partial x_1}dx_1 + \dfrac{\partial h}{\partial x_2}dx_2 + \dfrac{\partial h}{\partial x_3}dx_3 = 0$

- **제약조건 h가 있기 때문에 dx_1, dx_2, dx_3가 1차 독립이 아님.**

$$df + \lambda \cdot dh = 0$$

$$\left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right)dx_1 + \left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right)dx_2 + \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right)dx_3 = 0$$

λ값을 적절히 결정하여 첫번째 괄호 안의 값을 0으로 만들면 위 식은 dx_1에 무관하게 성립한다.

$$\left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right)dx_2 + \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right)dx_3 = 0$$

dx_2와 dx_3는 임의의 값이므로 위 식을 항상 만족하기 위해서는 괄호 안의 식이 0이 되어야 한다.
(dx_2, dx_3 를 임의의 값으로 택해도, dx_1이 dx_2, dx_3에 따라 결정되어, 식②를 만족)

따라서 다음을 만족하는 값 λ, x_1, x_2, x_3을 찾으면 그 점이 상점이 된다. ⇨ **미지수: 4개** (x_1, x_2, x_3, λ)

$$\therefore \left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right) = 0,\quad \left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right) = 0,\quad \left(\frac{\partial f}{\partial x_3} + \lambda\frac{\partial h}{\partial x_3}\right) = 0,\quad h(x_1, x_2, x_3) = 0$$

⇨ **식: 4개** ⇨ **유일해 존재**

제약조건이 있는 경우의 함수와 상점(Stationary Point)

Given: $minimize\ f(x_1, x_2, x_3)$

$h(x_1, x_2, x_3) = 0$

Find: **상점** (x_1^*, x_2^*, x_3^*)

다음을 만족하는 값 λ, x_1, x_2, x_3 **을 찾으면 그 점이 상점이 된다.**

$$\frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0, \quad \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0$$

$$\frac{\partial f}{\partial x_3} + \lambda \frac{\partial h}{\partial x_3} = 0, \quad h(x_1, x_2, x_3) = 0$$

제약조건이 있는 최적화 문제를 제약조건이 없는 최적화 문제로 변경

Lagrange **함수**(L)**를 정의 하고** L**의 상점을 구함**

$$L(x_1, x_2, x_3, \lambda) = f(x_1, x_2, x_3) + \lambda h(x_1, x_2, x_3)$$

$$\nabla L(x_1, x_2, x_3, \lambda) = 0$$

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} = 0 \qquad \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} = 0$$

$$\frac{\partial L}{\partial x_3} = \frac{\partial f}{\partial x_3} + \lambda \frac{\partial h}{\partial x_3} = 0 \qquad \frac{\partial L}{\partial \lambda} = h(x_1, x_2, x_3) = 0$$

λ : Lagrange Multiplier

L : Lagrange Function

[요약] 제약조건이 있는 경우의 함수와 상점

- 제약 최적화 문제와 그 해법 - Lagrange Multiplier를 이용

최적화 문제

Minimize $f(x_1, x_2, x_3)$ ···· ①

Subject to $h_1(x_1, x_2, x_3) = 0$ ···· ②

$h_2(x_1, x_2, x_3) = 0$ ···· ③

함수 f 값이 최소가 되기 위한 필요조건: 식 ①’

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 = 0 \cdots ①’$$

$$h_1(x_1, x_2, x_3) = 0 \cdots ②$$

$$h_2(x_1, x_2, x_3) = 0 \cdots ③$$

식 h_1, h_2로 인하여 dx_1, dx_2, dx_3는 서로 독립이 아니다.

dx_1, dx_2, dx_3 의 관계를 알기 위해 식②, ③를 변형

식 ②, ③ 으로부터 식 ②’, ③’를 유도

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \frac{\partial f}{\partial x_3}dx_3 = 0 \cdots ①’$$

$$dh_1 = \frac{\partial h_1}{\partial x_1}dx_1 + \frac{\partial h_1}{\partial x_2}dx_2 + \frac{\partial h_1}{\partial x_3}dx_3 = 0 \cdots ②’$$

$$dh_2 = \frac{\partial h_2}{\partial x_1}dx_1 + \frac{\partial h_2}{\partial x_2}dx_2 + \frac{\partial h_2}{\partial x_3}dx_3 = 0 \cdots ③’$$

식 ①’의 dx_1, dx_2를 소거하기 위하여 식 ②’, ③’을 이용하여 다음 식을 유도

$$df + \lambda_1 dh_1 + \lambda_2 dh_2 = 0 \cdots$$ ①’에 새로운 변수 λ_1, λ_2를 도입하고 ②’, ③’을 대입 → 식④,⑤,⑥을 유도

$$\underbrace{\left(\frac{\partial F}{\partial x_1} + \lambda_1\frac{\partial f_1}{\partial x_1} + \lambda_2\frac{\partial f_2}{\partial x_1}\right)dx_1}_{=0\ (dx_1\text{를 소거하기 위함, 식④})} + \underbrace{\left(\frac{\partial F}{\partial x_2} + \lambda_1\frac{\partial f_1}{\partial x_2} + \lambda_2\frac{\partial f_2}{\partial x_2}\right)dx_2}_{=0\ (dx_2\text{를 소거하기 위함, 식⑤})} + \underbrace{\left(\frac{\partial F}{\partial x_3} + \lambda_1\frac{\partial f_1}{\partial x_3} + \lambda_2\frac{\partial f_2}{\partial x_3}\right)dx_3}_{=0\ (dx_3\text{는 독립변수 이므로, 식⑥})} = 0$$

$$dh_1 = \frac{\partial h_1}{\partial x_1}dx_1 + \frac{\partial h_1}{\partial x_2}dx_2 + \frac{\partial h_1}{\partial x_3}dx_3 = 0 \cdots ②’$$

$$dh_2 = \frac{\partial h_2}{\partial x_1}dx_1 + \frac{\partial h_2}{\partial x_2}dx_2 + \frac{\partial h_2}{\partial x_3}dx_3 = 0 \cdots ③’$$

[요약] 제약조건이 있는 경우의 함수와 상점

- 제약 최적화 문제와 그 해법 - Lagrange Multiplier를 이용

최적화 문제

Minimize $f(x_1, x_2, x_3)$ …. ①

Subject to $h_1(x_1, x_2, x_3) = 0$ …. ②

$h_2(x_1, x_2, x_3) = 0$ …. ③

$$\underbrace{\left(\frac{\partial F}{\partial x_1} + \lambda_1 \frac{\partial f_1}{\partial x_1} + \lambda_2 \frac{\partial f_2}{\partial x_1}\right)}_{=0\ (dx_1\text{를 소거하기 위함, 식④})} dx_1 + \underbrace{\left(\frac{\partial F}{\partial x_2} + \lambda_1 \frac{\partial f_1}{\partial x_2} + \lambda_2 \frac{\partial f_2}{\partial x_2}\right)}_{=0\ (dx_2\text{를 소거하기 위함, 식⑤})} dx_2 + \underbrace{\left(\frac{\partial F}{\partial x_3} + \lambda_1 \frac{\partial f_1}{\partial x_3} + \lambda_2 \frac{\partial f_2}{\partial x_3}\right)}_{=0\ (dx_3\text{는 독립변수 이므로, 식⑥})} dx_3 = 0$$

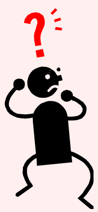
$$dh_1 = \frac{\partial h_1}{\partial x_1} dx_1 + \frac{\partial h_1}{\partial x_2} dx_2 + \frac{\partial h_1}{\partial x_3} dx_3 = 0 \quad \cdots \ ②'$$

$$dh_2 = \frac{\partial h_2}{\partial x_1} dx_1 + \frac{\partial h_2}{\partial x_2} dx_2 + \frac{\partial h_2}{\partial x_3} dx_3 = 0 \quad \cdots \ ③'$$

함수 f, h_1, h_2는 이미 주어진 함수이기 때문에 식 ④, ⑤, ⑥은 x_1, x_2, x_3의 (비)선형 대수 방정식 형태이다.

식 ④, ⑤, ⑥은 식 ①로 부터 유도 되었다.(변수 λ_1, λ_2 가 추가되어 식의 개수가 3개가 됨)

변수가 총 5개($x_1, x_2, x_3, \lambda_1, \lambda_2$)이므로 식 ②, ③, ④, ⑤, ⑥을 사용하여 문제를 풀 수 있다.



1. 식 ①', ②', ③'은 미분 방정식이 아닌가요?

→ 만일 다음과 같은 문제라면 미분 방정식이 맞습니다.

- Given: $\frac{\partial F}{\partial x_1} dx_1 + \frac{\partial F}{\partial x_2} dx_2 + \frac{\partial F}{\partial x_3} dx_3 = 0$, $\frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 + \frac{\partial f_1}{\partial x_3} dx_3 = 0$, $\frac{\partial f_2}{\partial x_1} dx_1 + \frac{\partial f_2}{\partial x_2} dx_2 + \frac{\partial f_2}{\partial x_3} dx_3 = 0$, - Find: 함수 F, f_1, f_2

그런데 함수 F, f_1, f_2(식①, ②, ③)는 이미 주어진 것이고, 우리가 구하려고 하는 것은 x_1, x_2, x_3이기 때문에 미분 방정식이 아닙니다.

공학수학 Review

- Quadratic Forms

Advanced
Ship
Design
Automation
Laboratory

Quadratic Forms. Transformation to Principal Axes

Definition: a quadratic form Q in the components $x_1, \dots, x_n$ of a vector $\mathbf{x}$ is a sum of n^2 terms, namely,

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k$$

$$= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n$$

$$+ a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots +$$

$$+ a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2$$

$\mathbf{A} = [a_{jk}]$ is called of the *coefficient matrix* of the form. We may assume that $\mathbf{A}$ is *symmetric*, because we can take off-diagonal terms together in pairs and write the result as a sum of two equal terms.

Quadratic Forms. Transformation to Principal Axes

Let

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^2 + \boxed{4x_1x_2 + 6x_2x_1} + 2x_2^2$$

$$= 3x_1^2 + \boxed{10x_1x_2} + 2x_2^2$$

$$= 3x_1^2 + \boxed{5x_1x_2 + 5x_2x_1} + 2x_2^2$$

$$= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x} \qquad \mathbf{C} = [c_{jk}], \; c_{jk} = \frac{1}{2}(a_{jk} + a_{kj})$$

Principal Axes Theorem

Theorem 8.4.2 Symmetric Matrices

A symmetric matrix has an orthonormal basis of eigenvectors for R^n.

(Quadratic Forms) $$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k$$

By the Theorem 8.4.2 the *symmetric* coefficient matrix $\mathbf{A}$ has an *orthonormal basis of eigenvectors* $\mathbf{x}_1, \dots, \mathbf{x}_n$. Let $\mathbf{X}$ be

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix}$$

$\mathbf{X}$ is *orthogonal*, so that $\mathbf{X}^{-1}=\mathbf{X}^{\mathrm{T}}$, we obtain

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

$$\therefore \mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \mathbf{X}\mathbf{D}\mathbf{X}^{\mathrm{T}}$$

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{x}^{\mathrm{T}}(\mathbf{X}\mathbf{D}\mathbf{X}^{\mathrm{T}})\mathbf{x}$$

Symmetric: $\mathbf{A}^{\mathrm{T}}=\mathbf{A}$
Skew-symmetric: $\mathbf{A}^{\mathrm{T}}=-\mathbf{A}$
Orthogonal: $\mathbf{A}^{\mathrm{T}}=\mathbf{A}^{-1}$

Principal Axes Theorem

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{X}\mathbf{D}\mathbf{X}^{\mathrm{T}}\mathbf{x}$$

If we set $\mathbf{X}^{\mathrm{T}}\mathbf{x} = \mathbf{y}$, then, since $\mathbf{X}^{\mathrm{T}} = \mathbf{X}^{-1}$, we get

$$\mathbf{x} = \left(\mathbf{X}^{\mathrm{T}}\right)^{-1}\mathbf{y} = \mathbf{X}\mathbf{y}$$

Furthermore, we have

$$\mathbf{x}^{\mathrm{T}}\mathbf{X} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{x}\right)^{\mathrm{T}} = \mathbf{y}^{\mathrm{T}}$$

So Q becomes simply

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{X}\mathbf{D}\mathbf{X}^{\mathrm{T}}\mathbf{x} = \mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y}$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

Principal Axes Theorem

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \longleftarrow \mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

$$= \mathbf{x}^{\mathrm{T}}\mathbf{X}\mathbf{D}\mathbf{X}^{\mathrm{T}}\mathbf{x} \longleftarrow \mathbf{y} = \mathbf{X}^{-1}\mathbf{x} = \mathbf{X}^{\mathrm{T}}\mathbf{x}$$

$$= \mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y} \qquad \mathbf{y}^{\mathrm{T}} = (\mathbf{X}^{\mathrm{T}}\mathbf{x})^{\mathrm{T}} = \mathbf{x}\mathbf{X}^{\mathrm{T}}$$

Theorem 8.4.5 Principal Axes Theorem

The substitution $\mathbf{x} = \mathbf{X}\mathbf{y}$ transforms a quadratic form

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \qquad (a_{kj} = a_{jk})$$

to the *principal axes form* or *canonical form*

$$Q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

where $\lambda_1, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of the *symmetric* matrix $\mathbf{A}$, and $\mathbf{X}$ is an *orthogonal* matrix with corresponding eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$, respectively, as column vectors.

Principal Axes Theorem

Ex) Transformation to Principal Axes. Conic Sections

Find out what type of conic section the following quadratic form represents and transform it to principal axes. (Ex 8.2-1)

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128$$

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = \mathbf{x}^T\mathbf{A}\mathbf{x}$$

$$= [x_1 \quad x_2]\begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}$$

Characteristic Equation:

$$\begin{vmatrix} 17-\lambda & -15 \\ -15 & 17-\lambda \end{vmatrix} = 0$$

$$(17-\lambda)^2 - 15^2 = 0 \qquad \therefore \lambda_1 = 2,\ \lambda_2 = 32$$

Eigenvectors

$$\lambda_1 = 2, \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 32, \mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$Q = \mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{x} = \mathbf{y}^T\mathbf{D}\mathbf{y}$$

$$= [y_1, y_2]\begin{bmatrix} 2 & 0 \\ 0 & 32 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= 2y_1^2 + 32y_2^2 = 128$$

$$\therefore \frac{y_1^2}{8^2} + \frac{y_2^2}{2^2} = 1$$

Ex) Transformation to Principal Axes. Conic Sections

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128$$
$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$$
$$\lambda_1 = 2,\ \lambda_2 = 32$$

Principal Axes

$$\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}$$

1) $\lambda = \lambda_1 = 2$

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 15 & -15 \\ -15 & 15 \end{bmatrix}$$

$$15x_1 - 15x_2 = 0$$

From this we get normalized eigenvector $\mathbf{x}_1$.

$$\mathbf{x}_1 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{\mathrm{T}}$$

2) $\lambda = \lambda_2 = 32$

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} -15 & -15 \\ -15 & -15 \end{bmatrix}$$

$$-15x_1 - 15x_2 = 0$$

From this we get normalized eigenvector $\mathbf{x}_2$.

$$\mathbf{x}_2 = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{\mathrm{T}}$$

Principal Axes Theorem

Ex) Transformation to Principal Axes. Conic Sections

$$\mathbf{x} = \mathbf{X}\mathbf{y}$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}$$

$$\mathbf{x}_1 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{\mathbf{T}}$$

$$\mathbf{x}_2 = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^{\mathbf{T}}$$

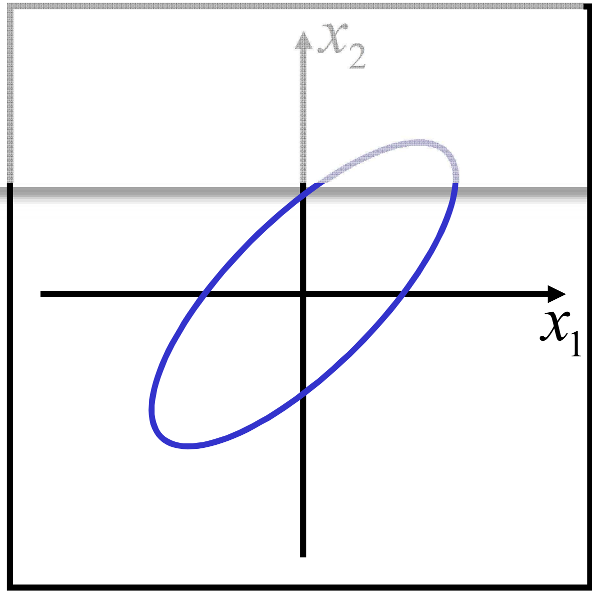
$$\therefore \mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

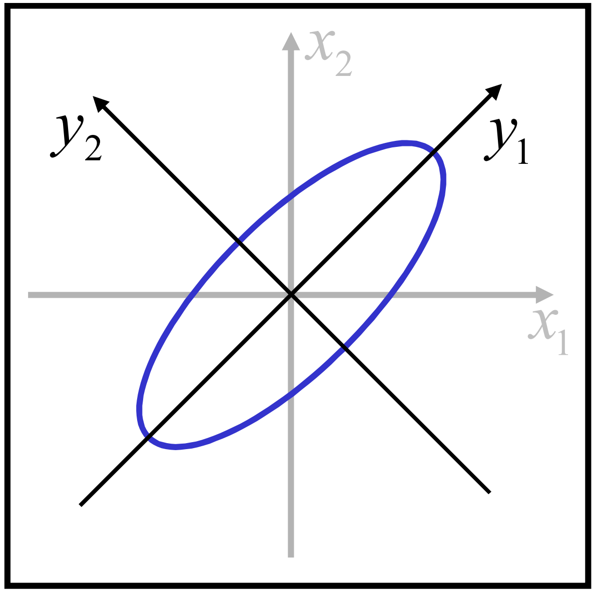
$$= \begin{bmatrix} \cos 45° & -\sin 45° \\ \sin 45° & \cos 45° \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128$$

$$\mathbf{x} = \mathbf{X}\mathbf{y}$$

$$\frac{y_1^2}{8^2} + \frac{y_2^2}{2^2} = 1$$





➔This means a 45° rotation (of principal axes. (See example 8.2-1.)

Quadratic form(Definiteness)

A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ and its (symmetric!) matrix A are called

(a) positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$,

(b) negative definite if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq 0$,

(c) indefinite if Q(x) takes both positive and negative values.

Quadratic form(Definiteness)

A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ and its (symmetric!) matrix A are called

(a) positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$,

(b) negative definite if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq 0$,

(c) indefinite if Q(x) takes both positive and negative values.

A necessary and sufficient condition for positive definiteness is that all the "principal minors" are positive, that is,

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \cdots \quad , \quad \det \mathbf{A} > 0$$

Show that the form in Prob. 23 is positive definite, whereas that in Prob. 19 is indefinite.

Quadratic form(Definiteness)

A necessary and sufficient condition for positive definiteness is that all the "principal minors" are positive, that is,

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \cdots, \det \mathbf{A} > 0$$

$$\mathbf{A} = \begin{bmatrix} 4 & \sqrt{3} \\ \sqrt{3} & 2 \end{bmatrix}$$

$$a_{11} = 4 > 0$$

$$\begin{vmatrix} 4 & \sqrt{3} \\ \sqrt{3} & 2 \end{vmatrix} = 4 \cdot 2 - \left(\sqrt{3}\right)^2 = 5 > 0$$

➔ **positive definite**

$$\mathbf{A} = \begin{bmatrix} 1 & 12 \\ 12 & -6 \end{bmatrix}$$

$$a_{11} = 1 > 0$$

$$\begin{vmatrix} 1 & 12 \\ 12 & -6 \end{vmatrix} = -6 - 12^2 = -150 < 0$$

➔ **indefinite**

Quadratic form(Definiteness)

the eigenvalues of $\mathbf{A}$ are

(a) positive definite: all positive

(b) negative definite: all negative

(c) indefinite: both positive and negative

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}, \quad \mathbf{x} = \mathbf{X}\mathbf{y}$$

$$= \mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \quad \textbf{(1)}$$

Because y = X^{-1}x, if x≠0, then y≠0.

From equation (1),

If all eigenvalues are positive, $Q(x)$ is positive.

If all eigenvalues are negative, $Q(x)$ is negative.