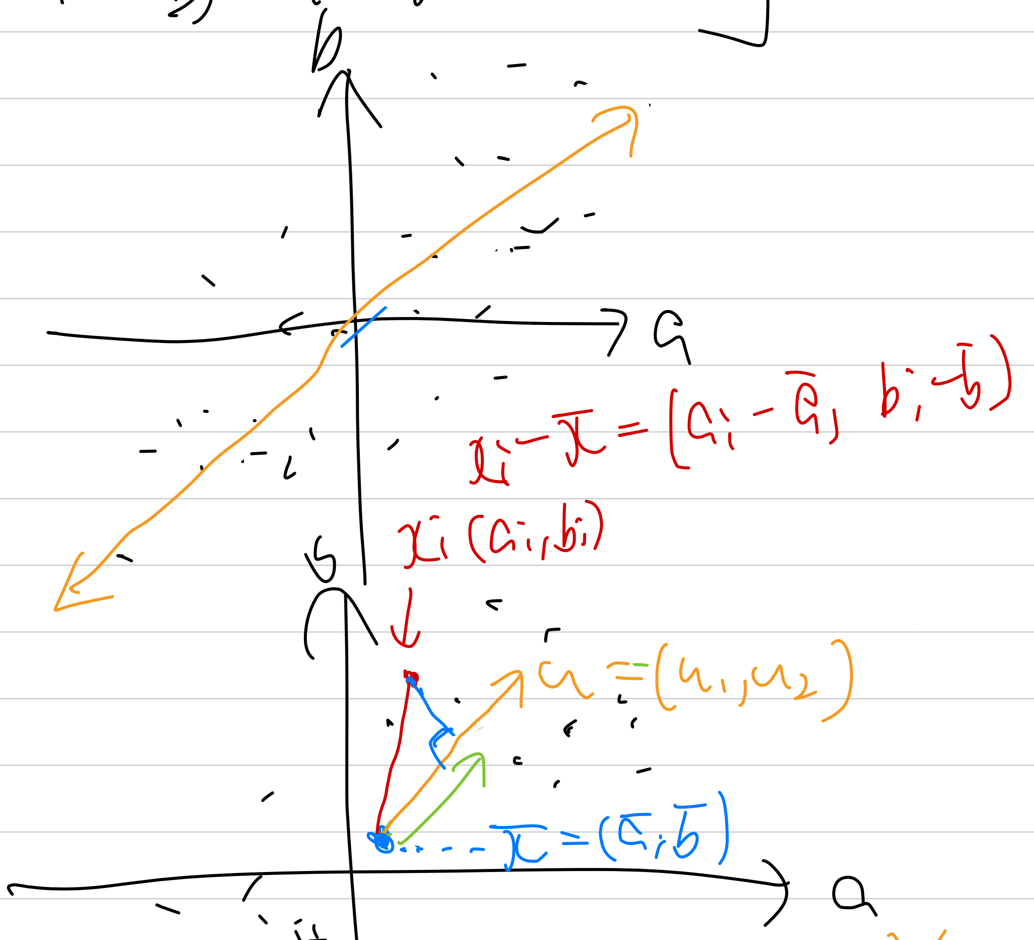
7/22

主成分とは
「2次元以上のサンプルが与えられた際に、
分散を最大にする方向である
」



オレンジのベクトルの方向が
主成分

分散を最大化する方向
||
つまり、「オレンジのベクトルが分散共分散行列の固有ベクトルと等しいこと」がわかればよい

総数nのi番目のサンプル：$x_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}$

サンプルの平均ベクトル：$\bar{x} = \begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix}$

単位ベクトル $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ を求めたい。

$x_i - \bar{x}$ から u の正射影ベクトルは

$$(x_i - \bar{x}) \cdot u \frac{u}{\|u\|}$$

正射影の長さは

$$(x_i - \bar{x}) \cdot u = (a_i - \bar{a})u_1 + (b_i - \bar{b})u_2 \cdots ①$$

$$= \begin{bmatrix} a_i - \bar{a} & b_i - \bar{b} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$= (x_i - \bar{x})^T u$$

x_i を u で回転させたもの
の分散を最大化したい

$x_i^T u$ の分散は

$$\frac{1}{n}\sum_{i=1}^{n}(x_i^T u - \bar{x}^T u)^2 = \frac{1}{n}\sum_{i=1}^{n}\left((x_i^T - \bar{x}^T)\cdot u\right)^2 \leftarrow ①$$

$$= \frac{1}{n}\sum_{i=1}^{n}\{(a_i-\bar{a})u_1 + (b_i-\bar{b})u_2\}^2$$

$$= \frac{1}{n}\underbrace{\left[(a_1-\bar{a})u_1+(b_1-\bar{b})u_2 \cdots (a_n-\bar{a})u_1+(b_n-\bar{b})u_2\right]}_{1\times n}\underbrace{\begin{bmatrix}(a_1-\bar{a})u_1+(b_1-\bar{b})u_2\\ \vdots \\ (a_n-\bar{a})u_1+(b_n-\bar{b})u_2\end{bmatrix}}_{n\times 1}$$

$$= \frac{1}{n}\underbrace{\begin{bmatrix}u_1 & u_2\end{bmatrix}}_{1\times 2}\underbrace{\begin{bmatrix}a_1-\bar{a} & \cdots & a_n-\bar{a}\\ b_1-\bar{b} & \cdots & b_n-\bar{b}\end{bmatrix}}_{2\times n}\underbrace{\begin{bmatrix}a_1-\bar{a} & b_1-\bar{b}\\ \vdots & \vdots\\ a_n-\bar{a} & b_n-\bar{b}\end{bmatrix}}_{n\times 2}\underbrace{\begin{bmatrix}u_1\\ u_2\end{bmatrix}}_{2\times 1}$$

$$= \begin{bmatrix}u_1 & u_2\end{bmatrix}\underbrace{\begin{bmatrix}\frac{1}{n}\sum_{i=1}^{n}(a_i-\bar{a})^2 & \frac{1}{n}\sum_{i=1}^{n}(a_i-\bar{a})(b_i-\bar{b})\\ \frac{1}{n}\sum_{i=1}^{n}(b_i-\bar{b})(a_i-\bar{a}) & \frac{1}{n}\sum_{i=1}^{n}(b_i-\bar{b})^2\end{bmatrix}}_{S\text{ は }2\times 2}\begin{bmatrix}u_1\\ u_2\end{bmatrix} \cdots ②$$

$$= u^T S u$$

S は $x_i = \begin{pmatrix} g_i \\ b_i \end{pmatrix}$ の分散共分散行列

$$\max_{u} \; u^T S u \qquad \text{s.t.} \; \|u\| = 1 \Leftrightarrow u^T u = 1$$

$$\Leftrightarrow u_1^2 + u_2^2 = 1$$

ラグランジュの未定乗数法により

$$\max_{u} \; L := u^T S u + \lambda(1 - u^T u)$$

$$\frac{\partial L}{\partial u} = 0 \quad \text{より、}$$

$$2Su - 2\lambda u = 0$$

$$Su = \lambda u \quad \cdots ③$$

これは分散共分散行列 S の $\underbrace{\text{固有値}}_{\lambda}$、$\underbrace{\text{固有ベクトル}}_{u}$ の定義式に等しい。

最適化された特徴量は $x_i^T u$

教科書では
直交化された特徴量は

$P = \underbrace{Z}_{Xの標準化}W \neq ZW$ となっている

X: $(t \times n)$ 行列

Z: $(t \times n)$ 行列となっているが

この例では

$x_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}$　　$\underbrace{X = \begin{bmatrix} a_1 & \cdots\cdots & a_t \\ b_1 & & b_t \end{bmatrix}}_{(n \times t)行列}$

と逆になっているから注意。