PCA

PRINCIPAL COMPONENT ANALYSIS

RECORDAR É VIVER...

PRODUTO INTERNO DE 2 VETORES $\underset{\sim}{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix}$ e $\underset{\sim}{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_p \end{pmatrix}$

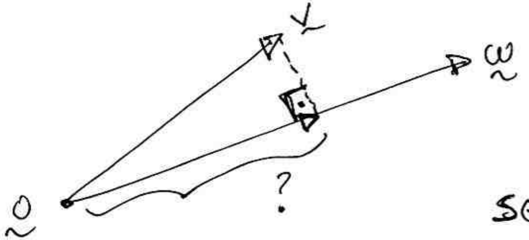
- $\langle \underset{\sim}{v}, \underset{\sim}{w} \rangle = \sum_{i=1}^{p} w_i v_i = (v_1 \cdots v_p) \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_p \end{pmatrix} = \underset{\sim}{v}^t \cdot \underset{\sim}{w}$

- $\underset{\sim}{v}^t \cdot \underset{\sim}{w} = \underset{\sim}{w}^t \cdot \underset{\sim}{v}$

- Como $\underset{\sim}{v}^t \cdot \underset{\sim}{w}$ é um NÚMERO REAL então

$$\left(\underset{\sim}{v}^t \cdot \underset{\sim}{w}\right)^2 = \left(\underset{\sim}{v}^t \cdot w\right)\left(\underset{\sim}{v}^t \cdot \underset{\sim}{w}\right) = \left(v^t \cdot \underset{\sim}{w}\right)\left(\underset{\sim}{w}^t \cdot \underset{\sim}{v}\right) = \underset{\sim}{v}^t \cdot \left(\underset{\sim}{w}\underset{\sim}{w}^t\right) \cdot \underset{\sim}{v}$$

- PROJEÇÃO ORTOGONAL ($\perp$) DE $\underset{\sim}{v}$ EM $\underset{\sim}{w}$:



$$? = \left(\underset{\sim}{v}^t \underset{\sim}{w}\right) \cdot \frac{\underset{\sim}{w}}{\|\underset{\sim}{w}\|^2}$$

SE $\|\underset{\sim}{w}\| = 1$ ENTÃO PROJEÇÃO $= \left(\underset{\sim}{v}^t \cdot \underset{\sim}{w}\right) \underset{\sim}{w}$

COVARIÂNCIA

$$Cov(Y_i, Y_j) = \underbrace{\mathbb{E}\left((Y_i - \mu_i)\cdot(Y_j - \mu_j)\right)} = \rho_{ij}\, \sigma_i\, \sigma_j$$

MATRIZ DE COVARIÂNCIA

$$\underset{p\times p}{\Sigma} = \mathbb{V}(\underset{\sim}{Y}) = \begin{bmatrix} \mathbb{V}(Y_1) & Cov(Y_1\ Y_2) & Cov(Y_1\ Y_3) & \cdots & Cov(Y_1\ Y_p) \\ Cov(Y_2\ Y_1) & \mathbb{V}(Y_2) & Cov(Y_2, Y_3) & \cdots & Cov(Y_2, Y_p) \\ \vdots & \vdots & \vdots & & \vdots \\ & & & & \mathbb{V}(Y_p) \\ Cov(Y_p, Y_1) & Cov(Y_p, Y_2) & Cov(Y_p, Y_3) & \cdots & \end{bmatrix}$$

Verifique que vale a igualdade abaixo (exercício):

$$\Sigma = \mathbb{E}\left[(\underset{\sim}{Y}-\underset{\sim}{\mu})(\underset{\sim}{Y}-\underset{\sim}{\mu})^t\right] = \mathbb{E}\left[\begin{pmatrix} Y_1-\mu_1 \\ \vdots \\ Y_p-\mu_p \end{pmatrix}(Y_1-\mu_1, \ldots, Y_p-\mu_p)\right] = \mathbb{E}\begin{bmatrix} (Y_1-\mu_1)^2 & (Y_1-\mu_1)(Y_2-\mu_2) & \cdots \\ (Y_2-\mu_2)(Y_1-\mu_1) & (Y_2-\mu_2)^2 & \cdots \\ \vdots & \vdots & \\ (Y_p-\mu_1)(Y_1-\mu_1) & (Y_p-\mu_p)(Y_2-\mu_2) & \cdots \end{bmatrix}$$

- $\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix} \sim N_p \left(\underset{p \times 1}{\underset{\sim}{\mu}} , \underset{p \times p}{\Sigma} \right)$

- CONTORNOS DE DENSIDADE CONSTANTE SÃO ELIPSÓIDES EM y DEFINIDOS POR

$$(\underset{\sim}{y} - \underset{\sim}{\mu})' \Sigma^{-1} (\underset{\sim}{y} - \underset{\sim}{\mu}) = c > 0$$

- ESTES ELIPSÓIDES SÃO CENTRADOS EM $\underset{\sim}{\mu}$ E TÊM EIXOS $\pm c \sqrt{\lambda_i}\, \underset{\sim}{v}_i$ ONDE $\Sigma \underset{\sim}{v}_i = \lambda_i \underset{\sim}{v}_i$ para $i = 1, 2, \ldots, p$

- ORDENAMOS OS AUTOVALORES DE FORMA QUE

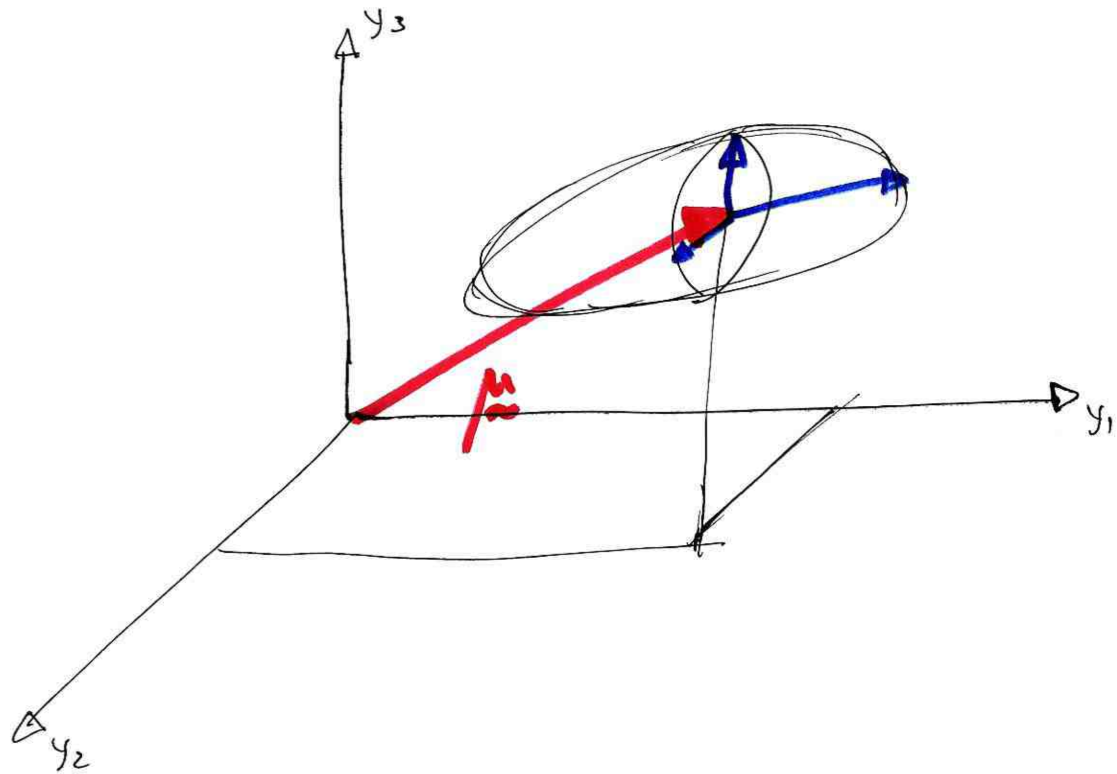
$$\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3 \geqslant \cdots \geqslant \lambda_p > 0$$

- CHAMAMOS $\underset{\sim}{v}_1$ DE PRIMEIRO AUTOVETOR OU 1º COMPONENTE PRINCIPAL

$\underset{\sim}{v}_2$ " SEGUNDO " " 2º "

ETC.

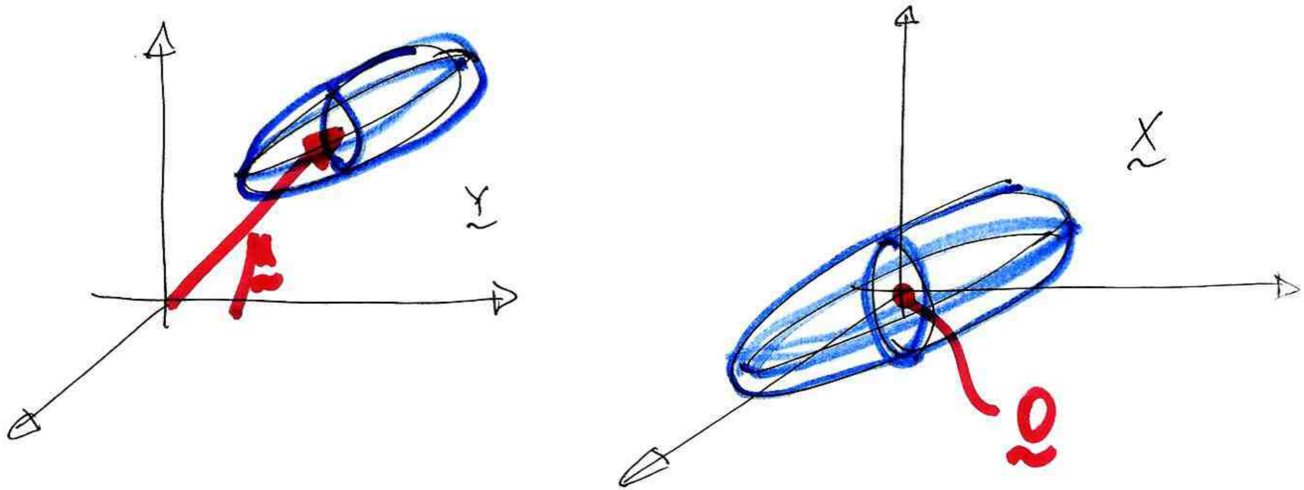
AMOSTRA DE $\underset{\sim}{Y} \sim N_3\left(\begin{pmatrix}\mu_1\\ \mu_2\\ \mu_3\end{pmatrix}, \Sigma\right)$



- Suponha que o menor autovalor λ_p seja ≈ 0 (5)
- Isto implica que o elipsóide e a nuvem de pontos será achatada.
- O que podemos fazer com isto?
- Isto implica que o vetor $\underset{\sim}{Y}$ vive aproximadamente num espaço de dimensão <u>MENOR</u> que sua dimensão p
- Podemos reduzir o nº de dimensões do vetor $\underset{\sim}{Y}$
- Esta redução está associada com os autovalores de Σ
- <u>Não tem</u> nenhuma relação com o vetor esperado $\underset{\sim}{\mu}$
- É mais fácil trabalhar com o vetor <u>$\underset{\sim}{Y}$ centrado em $\underset{\sim}{0}$</u> pois vamos trabalhar com <u>vetores e matrizes</u>.
- Defina $\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \underset{\sim}{Y} - \underset{\sim}{\mu} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix} - \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} Y_1 - \mu_1 \\ \vdots \\ Y_p - \mu_p \end{pmatrix}$

⊕ Veja que, se $\underset{\sim}{Y} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ então $\underset{\sim}{X} = \underset{\sim}{Y} - \underset{\sim}{\mu} \sim N_p(\underset{\sim}{0}, \Sigma)$ ⑥

$\Rightarrow$ MESMA Σ



⊕ SOMENTE DESLOCAMOS RIGIDAMENTE O ELIPSÓIDE DE FORMA QUE FIQUE CENTRADO NA ORIGEM.

⊕

* Pelo teorema espectral, podemos diagonalizar Σ (7)
usando seus autovetores:

$$\Sigma = P \Lambda P^t$$

onde $\underset{p\times p}{P} = \left[\underset{\sim}{v}_1 \mid \underset{\sim}{v}_2 \mid \cdots \mid \underset{\sim}{v}_p \right]$ é a matriz cujas colunas são os autovetores de Σ

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix}$$ é a matriz diagonal com os autovalores de Σ

e os autovetores são ortogonais e de comprimento 1 de forma que

$$\underset{\sim}{v}_i^t \cdot \underset{\sim}{v}_j = \delta_{ij} = \begin{cases} 1, \text{ se } i=j \\ 0, \text{ caso contrário} \end{cases}$$

(Ou seja, $P^t . P = I$)

- COMO OS ~~p~~ AUTOVETORES SÃO ORTOGONAIS, eles formam uma base do espaço $\mathbb{R}^p$ (8)

- O vetor aleatório $\underset{\sim}{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$ pode ser escrito como

$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} = c_1 \underset{\sim}{V}_1 + c_2 \underset{\sim}{V}_2 + \dots + c_p \underset{\sim}{V}_p$$

- Os valores de c_i são obtidos facilmente usando que os v_i's são $\perp$'s e têm comprimento 1. Por exemplo, para c_1:

$$\underset{\sim}{V}_1^t \cdot \underset{\sim}{X} = \underset{\sim}{V}_1^t \left(c_1 \underset{\sim}{V}_1 + c_2 \underset{\sim}{V}_2 + \dots + c_p \underset{\sim}{V}_p \right)$$

$$= c_1 \underbrace{\left(\underset{\sim}{V}_1^t \cdot \underset{\sim}{V}_1\right)}_{= \|\underset{\sim}{V}_1\|^2 = 1^2 = 1} + c_2 \underbrace{\left(\underset{\sim}{V}_1^t \underset{\sim}{V}_2\right)}_{= 0 \text{ pois } \perp} + \dots + c_p \underbrace{\left(\underset{\sim}{V}_1^t \underset{\sim}{V}_p\right)}_{= 0 \text{ pois } \perp} = c_1$$

Isto é, $c_1 = \underset{\sim}{V}_1^t \cdot \underset{\sim}{X}$

Do mesmo modo, para obter c_2

$$\underset{\sim}{V}_2' \cdot \underset{\sim}{X} = \underset{\sim}{V}_2' \left(c_1 \underset{\sim}{V}_1 + c_2 \underset{\sim}{V}_2 + \dots + c_p \underset{\sim}{V}_p \right)$$

$$= c_1 \underbrace{\underset{\sim}{V}_2' \underset{\sim}{V}_1}_{\substack{\| \\ 0}} + c_2 \underbrace{\underset{\sim}{V}_2' \underset{\sim}{V}_2}_{\substack{\| \\ 1}} + \dots + c_p \underbrace{\underset{\sim}{V}_2' \underset{\sim}{V}_p}_{\substack{\| \\ 0}} = c_2$$

Assim, a representação de $\underset{\sim}{X} = \underset{\sim}{Y} - \underset{\sim}{\mu}$ na base $\underset{\sim}{V}_1, \underset{\sim}{V}_2, \dots, \underset{\sim}{V}_p$ de autovetores de Σ é dada por

$$\underset{\sim}{X} = \underbrace{\left(\underset{\sim}{V}_1' \cdot \underset{\sim}{X}\right)}_{\in \mathbb{R}} \underset{\sim}{V}_1 + \underbrace{\left(\underset{\sim}{V}_2' \cdot \underset{\sim}{X}\right)}_{\in \mathbb{R}} \underset{\sim}{V}_2 + \dots + \underbrace{\left(\underset{\sim}{V}_p' \underset{\sim}{X}\right)}_{\in \mathbb{R}} \underset{\sim}{V}_p$$

- Suponha que os primeiros K autovalores sejam muito maiores que os últimos p-K
- Vamos supor que $\begin{cases} \lambda_{K+1} \approx 0 \\ \lambda_{K+2} \approx 0 \\ \vdots \\ \lambda_p \approx 0 \end{cases}$
- Vamos usar uma aproximaç X^* para X dada apenas pelas K primeiras coordenadas:

$$\underset{\sim}{X} = c_1 \begin{pmatrix} \underset{\sim}{V}_1 \end{pmatrix} + c_2 \begin{pmatrix} \underset{\sim}{V}_2 \end{pmatrix} + \ldots + c_K \begin{pmatrix} \underset{\sim}{V}_K \end{pmatrix} + c_{K+1} \begin{pmatrix} \underset{\sim}{V}_{K+1} \end{pmatrix} + \ldots + c_p \begin{pmatrix} \underset{\sim}{V}_p \end{pmatrix}$$

$$\approx \underbrace{c_1 \underset{\sim}{V}_1 + c_2 \underset{\sim}{V}_2 + \ldots + c_K \underset{\sim}{V}_K}_{= X^*} + \cancel{c_{K+1} \begin{pmatrix} \underset{\sim}{V}_{K+1} \end{pmatrix}} + \ldots + \cancel{c_p \begin{pmatrix} \underset{\sim}{V}_p \end{pmatrix}}$$

$$\boxed{\text{LEMBRE-SE: } c_i = \underset{\sim}{V}_i' \cdot \underset{\sim}{X}}$$

- Vamos mostrar agora que $\underset{\sim}{X} \approx \underset{\sim}{X}^*$!!
- AMBOS SÃO vetores do $\mathbb{R}^p$.

$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix} \qquad \underset{\sim}{X}^* = c_1 \underbrace{\begin{pmatrix} \underset{\sim}{v}_1 \end{pmatrix}}_{\in \mathbb{R}^p} + \dots + c_k \underbrace{\begin{pmatrix} \underset{\sim}{v}_k \end{pmatrix}}_{\in \mathbb{R}^p}$$

- Qual a vantagem então ??
- Suponha que $p = 20000$ e $k = 10$
- O vetor $\underset{\sim}{X}$ precisa de 20000 coordenadas.
- O vetor $\underset{\sim}{X}^*$ precisa de : 10 nº ($c_1, c_2, \dots, c_k$)
  e os k vetores $\underset{\sim}{v}_1, \dots, \underset{\sim}{v}_k$
- Para <u>várias instâncias</u> $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \dots, \underset{\sim}{X}_n$ os k vetores $\underset{\sim}{v}_1, \dots, \underset{\sim}{v}_k$ são os mesmos. Para diferenciá-los só precisamos de $c_1, \dots, c_k$ de cada um deles.

⊕ Isto ficará claro na aplicação de reconhecimento de faces (a seguir)

⊕ Agora vamos provar que $\underset{\sim}{X} \approx \underset{\sim}{X}^*$

⊕ $\underset{\sim}{X}$ e $\underset{\sim}{X}^*$ são vetores <u>aleatórios</u>

⊕ Para mostrar que, em geral, temos $\underset{\sim}{X} \approx X^*$, vamos tomar a distância (ao quadrado) entre eles

$$\underbrace{\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2}_{>0} \text{ e mostrar que } \mathbb{E}\left(\|X - X^*\|^2\right) \approx 0$$

~~PROVA:~~ $\cancel{\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \|c_1 \underset{\sim}{V}_1 + c_2 V_2 + \dots + c_k \underset{\sim}{V}_k + c_k V_{k+1} + \dots}$

⊕ **Prova:**

$$\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \left\| c_1 \underset{\sim}{V}_1 + \dots + c_k \underset{\sim}{V}_k + \underbrace{c_{k+1} \underset{\sim}{V}_{k+1} + \dots + c_p \underset{\sim}{V}_p}_{\text{SÓ VAI SOBRAR ISTO}} - \left(c_1 \underset{\sim}{V}_1 + \dots + c_p \underset{\sim}{V}_p\right) \right\|^2$$

Então

$$\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \|c_{k+1}\,\underset{\sim}{v}_{k+1} + \dots + c_p\,\underset{\sim}{v}_p\|^2$$

$$= c_{k+1}^2\,\|\underset{\sim}{v}_{k+1}\|^2 + \dots + c_p^2\,\|\underset{\sim}{v}_p\|^2 \quad \text{pois os } v_i\text{'s são } \perp\text{'s}$$

$$= c_{k+1}^2 + \dots + c_p^2 \quad \text{pois os } \|v_i\| = 1$$

$$= \left(\underset{\sim}{v}'_{k+1}\,\underset{\sim}{X}\right)^2 + \dots + \left(\underset{\sim}{v}'_p \cdot \underset{\sim}{X}\right)^2$$

$$= \left(\underset{\sim}{v}'_{k+1}\,\underset{\sim}{X}\right)\left(\underset{\sim}{X}^t\,\underset{\sim}{v}_{k+1}\right) + \dots + \left(\underset{\sim}{v}'_p\,\underset{\sim}{X}\right)\left(\underset{\sim}{X}^t\,\underset{\sim}{v}_p\right)$$

$$= \underset{\sim}{v}'_{k+1}\left(\underset{\sim}{X}\,\underset{\sim}{X}^t\right)\underset{\sim}{v}_{k+1} + \dots + v'_p\left(\underset{\sim}{X}\,\underset{\sim}{X}'\right)\underset{\sim}{v}_p$$

NOTE QUE

$$\underset{\sim}{X}\,\underset{\sim}{X}^t = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}(X_1 \dots X_p) = \begin{pmatrix} X_1^2 & X_1X_2 & \dots & X_1X_p \\ X_2X_1 & X_2^2 & \dots & X_2X_p \\ \vdots & \vdots & & \\ X_pX_1 & X_pX_2 & \dots & X_p^2 \end{pmatrix},$$

UMA MATRIZ $p \times p$

Encontramos que

$$\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \underset{\sim}{v}_{k+1}' \underset{\sim}{X} \underset{\sim}{X}' \underset{\sim}{v}_{k+1} + \dots + \underset{\sim}{v}_p' \underset{\sim}{X} \underset{\sim}{X}' \underset{\sim}{v}_p$$

Assim,

$$\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right) = \mathbb{E}\left(\underset{\sim}{v}_{k+1}' \underset{\sim}{X} \underset{\sim}{X}' \underset{\sim}{v}_{k+1} + \dots + v_p' \underset{\sim}{X} \underset{\sim}{X}' \underset{\sim}{v}_p\right)$$

$$= \underset{\sim}{v}_{k+1}' \underbrace{\mathbb{E}\left(\underset{\sim}{X} \underset{\sim}{X}'\right)}_{=\Sigma} \underset{\sim}{v}_{k+1} + \dots + \underset{\sim}{v}_p' \underbrace{\mathbb{E}\left(\underset{\sim}{X} \underset{\sim}{X}'\right)}_{=\Sigma} \underset{\sim}{v}_p$$

$$= \underset{\sim}{v}_{k+1}' \underbrace{\Sigma \underset{\sim}{v}_{k+1}}_{\lambda_{k+1} \underset{\sim}{v}_{k+1}} + \dots + \underset{\sim}{v}_p' \underbrace{\Sigma \underset{\sim}{v}_p}_{\lambda_p \underset{\sim}{v}_p}$$

$$= \lambda_{k+1} \underset{\sim}{v}_{k+1}' \underset{\sim}{v}_{k+1} + \dots + \lambda_p \underset{\sim}{v}_p' \underset{\sim}{v}_p$$

$$= \lambda_{k+1} \|\underset{\sim}{v}_{k+1}\|^2 + \dots + \lambda_p \|\underset{\sim}{v}_p\|^2 = \lambda_{k+1} + \dots + \lambda_p \approx 0$$

Portanto, encontramos

$$\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right) \approx 0$$

$\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2$ é uma v.a. positiva cujo valor esperado é aproximadamente $\underline{\underline{0}}$

Assim, esperamos $\|\underset{\sim}{X} - \underset{\sim}{X}^*\| = \sqrt{\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2} \approx 0$ TAMBÉM

MAS SE $\|\underset{\sim}{X} - \underset{\sim}{X}^*\| \approx 0$ então $\underset{\sim}{X} \approx \underset{\sim}{X}^*$

<u>Conclusão</u> Se os últimos autovalores forem próximos de zero podemos usar $\underset{\sim}{X}^*$ "=" $[c_1, \ldots, c_k]$ → representaç na base $\underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_k$

que teremos praticamente $\underset{\sim}{X}$.

- ISTO É O MELHOR QUE PODEMOS FAZER? (16)
- EXISTE OUTRO CONJUNTO ORTONORMAL

$\underset{\sim}{w}_1, \underset{\sim}{w}_2, \ldots, \underset{\sim}{w}_k$

TAL QUE A PROJEÇÃO $\perp$ $\underset{\sim}{X}^{**}$ DE $\underset{\sim}{X}$ NO ESPAÇO GERADO PELOS $\underset{\sim}{w}_i$'S SEJA MELHOR QUE A PROJEÇÃO $\underset{\sim}{X}^{*}$ NO ESPAÇO DOS K 1º AUTOVETORES $\underset{\sim}{V}_1, \underset{\sim}{V}_2, \ldots, \underset{\sim}{V}_k$?

- ISTO É, EXISTE $\underset{\sim}{w}_1, \underset{\sim}{w}_2, \ldots, \underset{\sim}{w}_k$ TAL QUE

$$\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^{**}\|^2\right) < \mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^{*}\|^2\right) ?$$

$\underset{\sim}{X}^{**}$ $\Downarrow$ projeção $\perp$ em $w_1, w_2, \ldots, w_k$

$\underset{\sim}{X}^{*}$ $\Downarrow$ projeção $\perp$ em $\underset{\sim}{V}_1, \ldots, \underset{\sim}{V}_k$ AUTOVETORES DE Σ

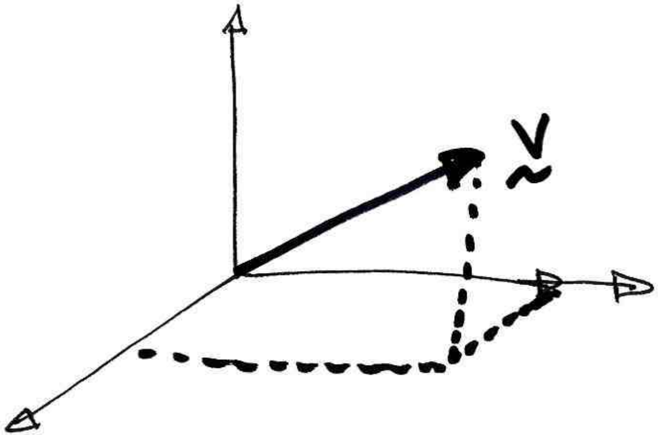
- Como você podia antecipar pelo estilo retórico da pergunta, a resposta é <u>NÃO</u>. 17

- O subespaço <u>de dimensão K</u> que minimiza o valor esperado do resíduo $E\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right)$ É O SUBESPAÇO FORMADO pelos 1º K autovetores de $\Sigma = \mathbb{V}ar(\underset{\sim}{X})$

- NENHUM OUTRO PODE REPRESENTAR $\underset{\sim}{X}$ TÃO BEM (EM K dimensões) quanto os K 1º autovetores de Σ.

- VAMOS PROVAR ISTO PARA $K=1$. MAS ANTES, UMAS FIGURAS...

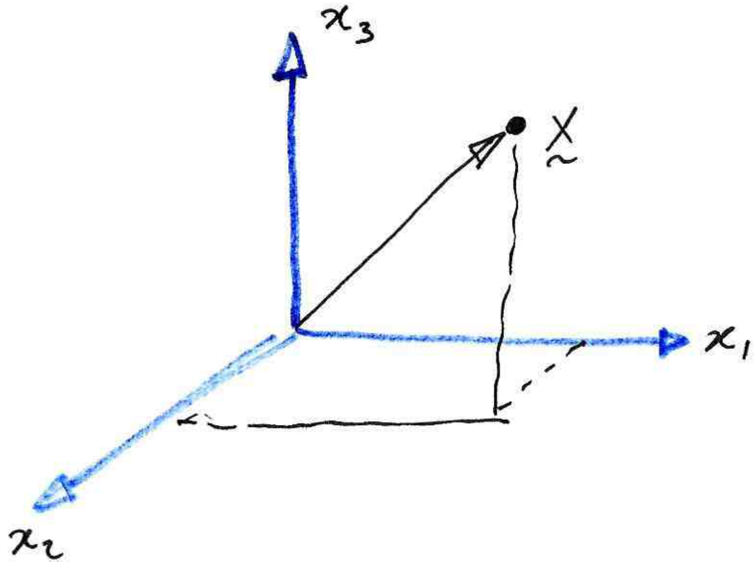
PROCURAMOS UMA LINHA PASSANDO PELA ORIGEM (SUBESPAÇO DE DIMENSÃO $K=1$) TAL QUE A DISTÂNCIA ALEATÓRIA $\|\underset{\sim}{X}-\underset{\sim}{X}^*\|^2$ ENTRE $\underset{\sim}{X}$ E A SUA PROJEÇÃO $\perp$ $\underset{\sim}{X}^*$ TENHA O VALOR ESPERADO MÍNIMO.



Linha $= \{c\underset{\sim}{v}$ onde $c \in \mathbb{R}\}$

$=$ MÚLTIPLOS DO VETOR $\underset{\sim}{v}$

Escolhendo diferentes $\underset{\sim}{v}$'s temos diferentes sub-espaços de dim $=1$.

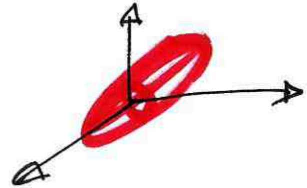


$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \text{ vive em } \mathbb{R}^3$$

$\underset{\sim}{X}$ é aleatório

$$\mathbb{E}(\underset{\sim}{X}) = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

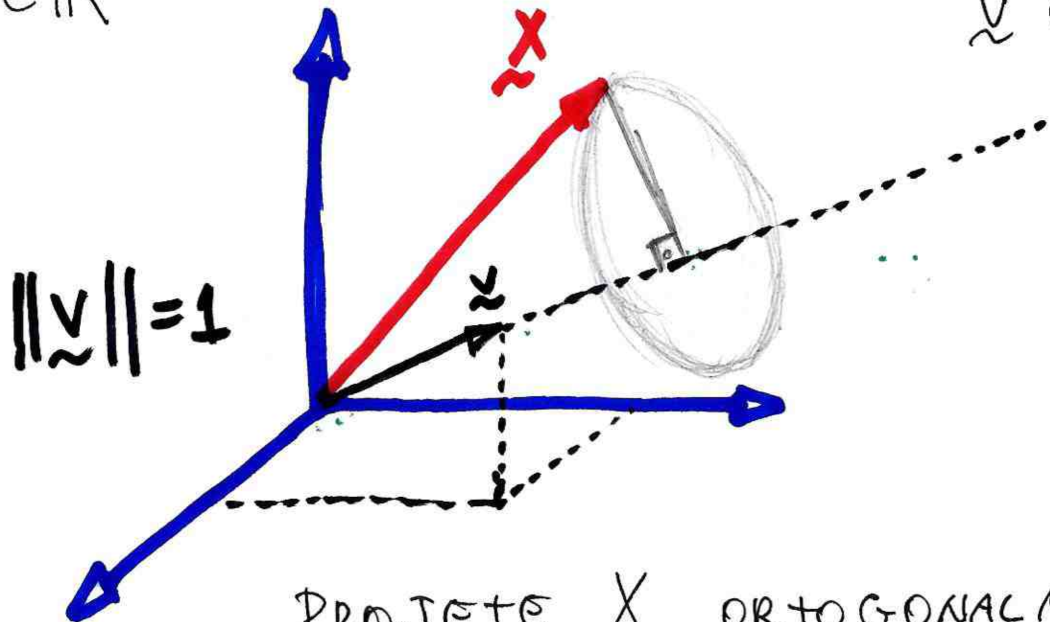
$\Longrightarrow$ SE AMOSTRAMOS $\underset{\sim}{X}$ VÁRIAS VEZES VEREMOS UMA NUVEM DE PONTOS CENTRADA NA ORIGEM NA FORMA DE UM ELIPSÓIDE



Fixe um $\underset{\sim}{v} \in \mathbb{R}^3$ qualquer e considere o vetor $\underset{\sim}{x} \in \mathbb{R}^3$

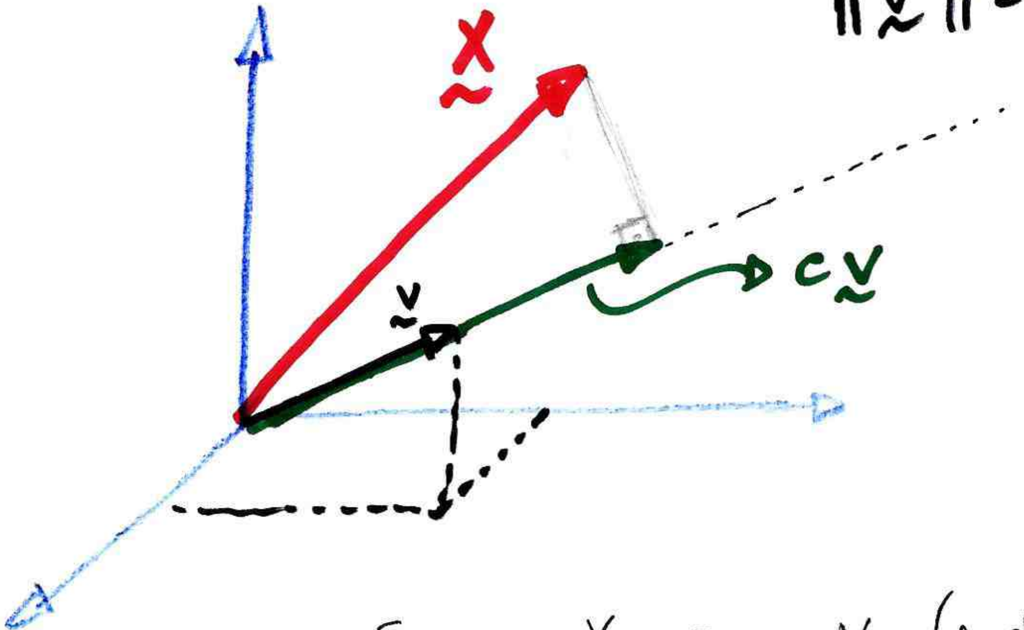
$\underset{\sim}{v} \in \mathbb{R}^3$ E COM

$\|\underset{\sim}{v}\| = 1$



PROJETE $\underset{\sim}{x}$ ORTOGONALMENTE EM $\underset{\sim}{v}$

$\|\underset{\sim}{V}\| = 1$



PROJEÇÃO DO $\underset{\sim}{X}$ EM $\underset{\sim}{V}$ (onde $\|\underset{\sim}{V}\|=1$)

é o vetor $\underset{\sim}{X}^* = c \cdot \underset{\sim}{V} = \underbrace{\left(\underset{\sim}{V}' \cdot \underset{\sim}{X}\right)}_{\text{MÚLTIPLO DO } \underset{\sim}{V}} * \underset{\sim}{V}$

Queremos encontrar $\underset{\sim}{v} \in \mathbb{R}^p$ com $\|\underset{\sim}{v}\| = 1$ tal que $\underset{\sim}{v}$ minimize $\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right) = \mathbb{E}\left(\|\underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v}\|^2\right)$ onde $\underset{\sim}{X}$ é vetor aleatório com $\mathbb{E}(\underset{\sim}{X}) = \underset{\sim}{0}$ e $\mathbb{V}(\underset{\sim}{X}) = \Sigma$

$$\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \|\underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v}\|^2 = \left\langle \underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v},\ \underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v} \right\rangle$$

$$= \left(\underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v}\right)' \cdot \left(\underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v}\right)$$

$$= \left(\underset{\sim}{X}' - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\cdot\underset{\sim}{v}'\right) \cdot \left(\underset{\sim}{X} - (\underset{\sim}{v}'\cdot\underset{\sim}{X})\underset{\sim}{v}\right)$$

$$= \underset{\sim}{X}'\underset{\sim}{X} - 2(\underset{\sim}{v}'\cdot\underset{\sim}{X})\,\underset{\sim}{X}'\cdot\underset{\sim}{v} + (\underset{\sim}{v}'\cdot\underset{\sim}{X})^2 \underbrace{\underset{\sim}{v}'\cdot\underset{\sim}{v}}_{=1 \text{ pois } \|\underset{\sim}{v}\| = 1}$$

Então $\|\underset{\sim}{X} - \underset{\sim}{X}^*\| = \underset{\sim}{X}'\underset{\sim}{X} - 2\,\underset{\sim}{v}'\underbrace{(\underset{\sim}{X}.\underset{\sim}{X}')}_{p\times p}\,\underset{\sim}{v} + (\underset{\sim}{v}'.\underset{\sim}{X})^2$

$$= \underset{\sim}{X}'\underset{\sim}{X} - 2\,\underset{\sim}{v}'(\underset{\sim}{X}.\underset{\sim}{X}')\underset{\sim}{v} + \underbrace{(\underset{\sim}{v}'\underset{\sim}{X}).(\underset{\sim}{v}'.\underset{\sim}{X})}$$

$$= (\underset{\sim}{v}'.\underset{\sim}{X})(\underset{\sim}{X}'.\underset{\sim}{v})$$

$$= \underbrace{\underset{\sim}{v}'(\underset{\sim}{X}.\underset{\sim}{X}')\underset{\sim}{v}}_{\text{veja o 2º termo}}$$

$$= \underset{\sim}{X}'\underset{\sim}{X} - \underset{\sim}{v}'(\underset{\sim}{X}.\underset{\sim}{X}')\underset{\sim}{v}$$

Queremos minimizar em $\underset{\sim}{v}$ a expressão

$$E(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2) = \underbrace{E(\underset{\sim}{X}'\underset{\sim}{X})}_{\text{NÃO DEPENDE DO } \underset{\sim}{v}.} - \underset{\sim}{v}'\underbrace{E(\underset{\sim}{X}.\underset{\sim}{X}')}_{\Sigma}\underset{\sim}{v} \Rightarrow$$

Minimizar $\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right)$ em $\underset{\sim}{v}$ é o mesmo
que maximizar $\underset{\sim}{v}^T \Sigma \underset{\sim}{v}$ com $\|\underset{\sim}{v}\| = 1$.

- Vamos provar que a soluç. é tomar $\underset{\sim}{v} = \underset{\sim}{v}_1$ = 1º autovetor de Σ
- Seja $\underset{\sim}{v} \in \mathbb{R}^p$ arbitrário com $\|\underset{\sim}{v}\|^2 = 1$.
- Podemos escrever $\underset{\sim}{v} = c_1 \underset{\sim}{v}_1 + c_2 \underset{\sim}{v}_2 + \dots + c_p \underset{\sim}{v}_p$ de forma única.
- Temos $1 = \|\underset{\sim}{v}\|^2 = c_1^2 \underbrace{\|\underset{\sim}{v}_1\|^2}_{=1} + c_2^2 \underbrace{\|\underset{\sim}{v}_2\|^2}_{=1} + \dots + c_p^2 \underbrace{\|\underset{\sim}{v}_p\|^2}_{=1}$

$$\Rightarrow \quad 1 = c_1^2 + c_1^2 + \dots + c_p^2$$

Pelo teorema espectral $\Sigma = P \Lambda P'$ onde

$$P = \left[\, v_1 \;\middle|\; v_2 \;\middle|\; \cdots \;\middle|\; v_p \,\right], \quad \Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \text{ e } P'P = PP' = I$$

Então

$$v' \Sigma v = \underbrace{v' P}_{= y'} \Lambda \underbrace{P' v}_{= y} = y' \Lambda y = \lambda_1 y_1^2 + \dots + \lambda_p y_p^2$$

$$\leq \lambda_1 y_1^2 + \lambda_1 y_2^2 + \dots + \lambda_1 y_p^2 \quad \text{pois } \lambda_1 = \max\{\lambda_i\}$$

$$= \lambda_1 (y_1^2 + \dots + y_p^2) = \lambda_1 \quad \text{pois}$$

$$y_1^2 + \dots + y_p^2 = \|y\|^2 = y' y = v' \underbrace{P P'}_{= I} \cancel{y} = v' v = \|v\| = 1$$

Portanto, $\underset{\sim}{v}' \Sigma \underset{\sim}{v} \leq \lambda_1 = $ 1º autovalor. (26)

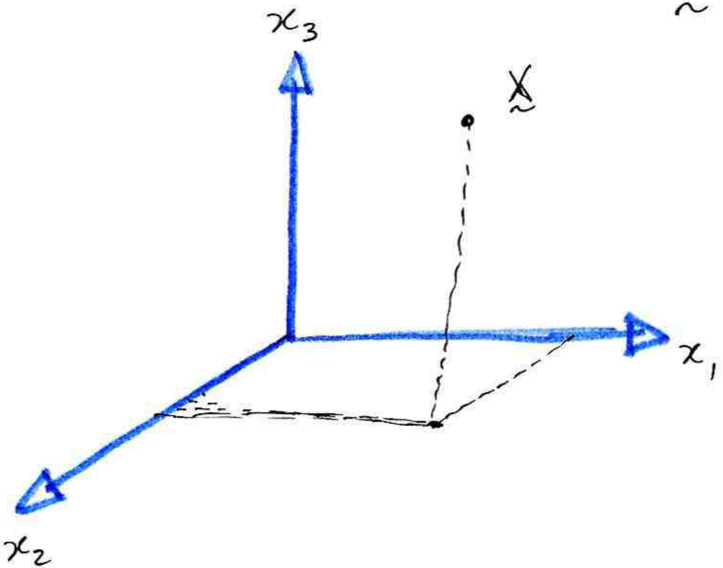
Por outro lado, se tomarmos $\underset{\sim}{v} = \underset{\sim}{v}_1 = $ 1º auto vetor

temos $\underset{\sim}{v}_1' \Sigma \underset{\sim}{v}_1 = \underset{\sim}{v}_1' (\lambda_1 \underset{\sim}{v}_1) = \lambda_1 \underset{\sim}{v}_1' \underset{\sim}{v}_1 = \lambda_1 \|\underset{\sim}{v}_1\|^2$
$= \lambda_1$

Assim, $\underset{\sim}{v} = \underset{\sim}{v}_1 = $ 1º autovetor atinge o ~~pto~~ máximo de $\underset{\sim}{v}' \Sigma \underset{\sim}{v}$ com $\|\underset{\sim}{v}\|^2 = 1$

Portanto, $\max\limits_{\underset{\sim}{v} \text{ t.q. } \|\underset{\sim}{v}\| = 1} \left(\underset{\sim}{v}' \Sigma \underset{\sim}{v} \right) = \lambda_1 = $ 1º autovalor

atingido quando $\underset{\sim}{v} = \underset{\sim}{v}_1 = $ 1º autovetor.



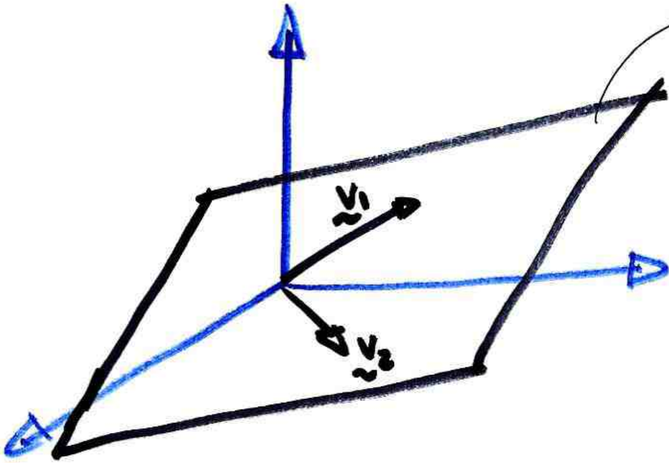
$\underset{\sim}{X}$ VIVE EM $\mathbb{R}^3$

$\underset{\sim}{X}$ É ALEATÓRIO

PROCURAMOS UM PLANO FIXO (K=2) PASSANDO PELA ORIGEM TAL QUE A DISTÂNCIA ALEATÓRIA $\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2$ ENTRE $\underset{\sim}{X}$ E A SUA PROJEÇÃO $\perp$ $\underset{\sim}{X}^*$ ~~SEJA~~ TENHA O VALOR ESPERADO MÍNIMO

28
27

PASSANDO para $K=2$:



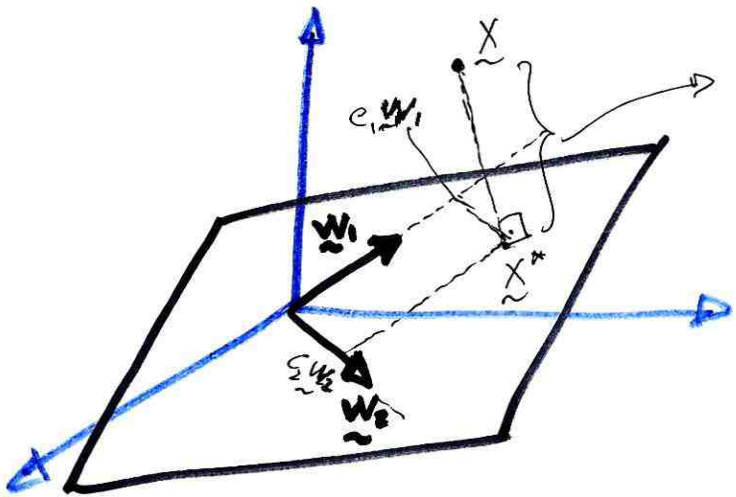
Plano é o conjunto de todas as combinações lineares de $\underset{\sim}{w}_1$ e $\underset{\sim}{w}_2$

$$\text{Plano} = \{ c_1 \underset{\sim}{w}_1 + c_2 \underset{\sim}{w}_2 \}$$

tomamos $\underset{\sim}{w}_1 \perp \underset{\sim}{w}_2$ e $\|\underset{\sim}{w}_1\| = \|\underset{\sim}{w}_2\| = 1$

Escolhendo diferentes bases, temos diferentes planos.

Fixe uma base qualquer e considere o vetor $\underset{\sim}{x} \in \mathbb{R}^3$



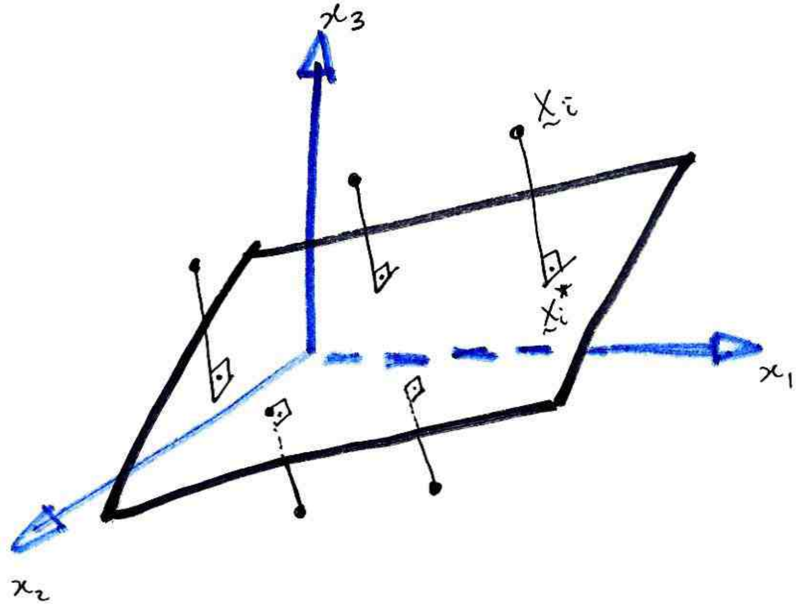
$\underset{\sim}{X} - \underset{\sim}{X}^*$

$\underset{\sim}{X}^* =$ Projeção $\perp$ de $\underset{\sim}{X}$ no plano formado por $\underset{\sim}{W_1}$ e $\underset{\sim}{W_2}$

$\underset{\sim}{X} - \underset{\sim}{X}^*$ É O VETOR DIFERENÇA

Queremos achar o plano tal que o comprimento <u>esperado</u> seja mínimo:

$$\mathbb{E}(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2) \text{ mínimo.}$$



AMOSTRA DE $\underline{\underline{M}}$ INSTÂNCIAS DO VETOR $\underset{\sim}{X}$ E SUAS RESPECTIVAS PROJEÇÕES ORTOGONAIS

PROCURAMOS O $\underline{\text{PLANO}}$ TAL QUE

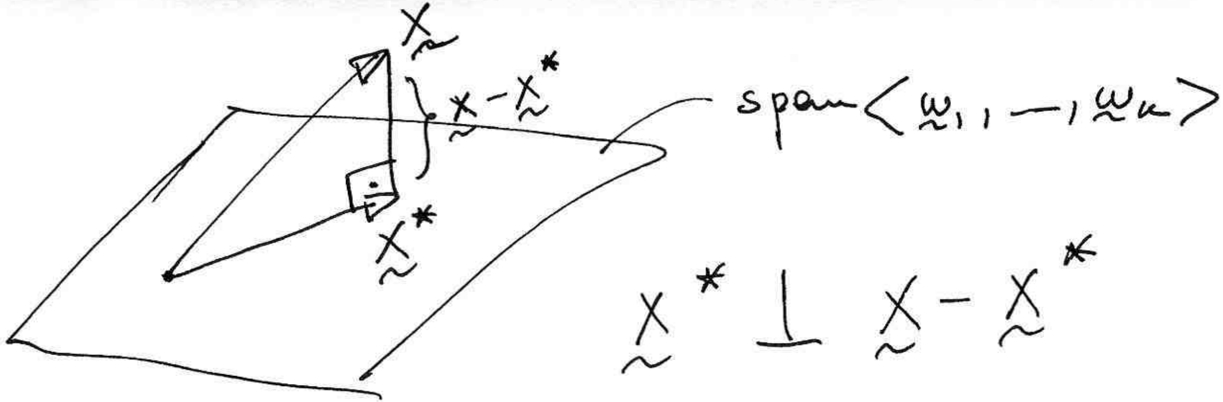
$$\underbrace{\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right)}_{\text{SEJA MÍNIMO}} \approx \underbrace{\frac{1}{M} \sum_{i=1}^{M} \|\underset{\sim}{x}_i - \underset{\sim}{x}_i^*\|^2}_{\text{MÉDIA ARITMÉTICA DAS DISTÂNCIAS}}$$

Generalização: Encontrar $\underset{\sim}{w}_1, \underset{\sim}{w}_2, \ldots, \underset{\sim}{w}_k \in \mathbb{R}^p$ (31)
que sejam ortonormais ($\perp$'s entre si e de comprimento 1)
tal que $\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right)$ seja mínimo onde

$\underset{\sim}{X}^* =$ projeção de $\underset{\sim}{X}$ em span$\langle \underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_k \rangle$.

Solução são os k primeiros autovetores de Σ.

Prova: Complete a ~~b~~ $\underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_k$ com $\underset{\sim}{w}_{k+1}, \ldots, \underset{\sim}{w}_p$
de forma a ter uma base ortonormal de $\mathbb{R}^p$.
A projeção $\perp$ de $\underset{\sim}{X} \in \mathbb{R}^p$ no span$\langle \underset{\sim}{w}_1, \ldots, \underset{\sim}{w}_k \rangle$ é
o vetor $\underset{\sim}{X}^* = a_1 \underset{\sim}{w}_1 + \ldots + a_k \underset{\sim}{w}_k$ tal que
$\underset{\sim}{X} - \underset{\sim}{X}^*$ é $\perp$ a $\underset{\sim}{X}^*$.



$$\underset{\sim}{X}^* \perp \underset{\sim}{X} - \underset{\sim}{X}^*$$

- Quem é $\underset{\sim}{X}^*$? Se $\underset{\sim}{X}^* = a_1 \underset{\sim}{w}_1 + \dots + a_k \underset{\sim}{w}_k$ quem são $a_1, a_2, \dots, a_k$?

- Sabemos que $\underset{\sim}{X}$ é escrito de forma única na base $\underset{\sim}{w}_1, \dots, \underset{\sim}{w}_p$ de $\mathbb{R}^p$: $\underset{\sim}{X} = c_1 \underset{\sim}{w}_1 + c_2 \underset{\sim}{w}_2 + \dots + c_p \underset{\sim}{w}_p$

com $c_i = \underset{\sim}{w}_i' \cdot \underset{\sim}{X}_i$

<u>Afirmação:</u> a projeção $\perp$ X^* é $\underset{\sim}{X}^* = c_1 \underset{\sim}{w}_1 + \dots + c_k \underset{\sim}{w}_k$

Isto é, basta tomar $a_i = c_i$

A prova é muito simples pois, como

$\underset{\sim}{X} - \underset{\sim}{X}^* = c_{k+1}\,\underset{\sim}{w}_{k+1} + \dots + c_p\,\underset{\sim}{w}_p$ então

$\underset{\sim}{X} - \underset{\sim}{X}^*$ é $\perp$ a $\underset{\sim}{X}^* = c_1\,\underset{\sim}{w}_1 + \dots + c_p\,\underset{\sim}{w}_p$

pois $\underset{\sim}{X}^* \in \text{span}\langle \underset{\sim}{w}_1, \dots, \underset{\sim}{w}_k \rangle$

$\underset{\sim}{X} - \underset{\sim}{X}^* \in \text{span}\langle w_{k+1}, \dots, \underset{\sim}{w}_p \rangle$

sub-espaços $\perp$'s entre si

Assim, a projeção $\perp$ de $\underset{\sim}{X}$ no subespaço $\text{span}\langle \underset{\sim}{w}_1, \dots, \underset{\sim}{w}_k \rangle$ é $\underset{\sim}{X}^* = c_1\,\underset{\sim}{w}_1 + \dots + c_k\,\underset{\sim}{w}_k$ onde

$$c_i = \underset{\sim}{w}_i' \cdot \underset{\sim}{X}$$

Queremos achar os vetores $\underset{\sim}{w}_1, \dots, \underset{\sim}{w}_k$ que minimizem o resíduo $\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right) \longrightarrow$ mínimo

Mas $\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \left(\underset{\sim}{X} - \underset{\sim}{X}^*\right)'\left(\underset{\sim}{X} - \underset{\sim}{X}^*\right) =$

$= \underset{\sim}{X}'\underset{\sim}{X} - 2\underset{\sim}{X}'\underset{\sim}{X}^* + \underset{\sim}{X}^{*\prime}\underset{\sim}{X}^*$

Temos $\underset{\sim}{X}'\underset{\sim}{X}^* = \left(c_1\underset{\sim}{\omega}_1' + c_2\underset{\sim}{\omega}_2' + \ldots + c_p\underset{\sim}{\omega}_p'\right)\cdot\left(c_1\underset{\sim}{\omega}_1 + \ldots + c_k\underset{\sim}{\omega}_k\right)$

$$= c_1^2\underbrace{\underset{\sim}{\omega}_1'\underset{\sim}{\omega}_1}_{=1} + c_2c_1\underbrace{\underset{\sim}{\omega}_1'\omega_2}_{=0} + \ldots + c_kc_1\underbrace{\underset{\sim}{\omega}_1'\underset{\sim}{\omega}_k}_{=0} +$$

$$+ c_1c_2\underbrace{\underset{\sim}{\omega}_2'\underset{\sim}{\omega}_1}_{=0} + c_2^2\underbrace{\underset{\sim}{\omega}_2'\omega_2}_{=1} + c_2c_3\underbrace{\underset{\sim}{\omega}_2'\omega_3}_{=0} + \ldots + c_2c_k\underbrace{\underset{\sim}{\omega}_2'\underset{\sim}{\omega}_k}_{=0} +$$

$$+ \ldots + c_pc_1\underbrace{\underset{\sim}{\omega}_p'\underset{\sim}{\omega}_1}_{=0} + c_pc_2\underbrace{\underset{\sim}{\omega}_p'\underset{\sim}{\omega}_2}_{=0} + \ldots + c_pc_k\underbrace{\underset{\sim}{\omega}_p'\underset{\sim}{\omega}_k}_{=0}$$

Isto é, $\underset{\sim}{X}'\underset{\sim}{X}^* = c_1^2 + \ldots + c_k^2$

Do mesmo modo, encontramos ~~$\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^* =$~~

$$\underset{\sim}{X}^{*\prime}\underset{\sim}{X}^* = c_1^2 + \ldots + c_k^2$$

Assim $\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2 = \underset{\sim}{X}'\underset{\sim}{X} - (c_1^2 + \ldots + c_k^2)$

e portanto

$$\mathbb{E}\left(\|\underset{\sim}{X} - \underset{\sim}{X}^*\|^2\right) = \underbrace{\mathbb{E}(\underset{\sim}{X}'\underset{\sim}{X})}_{\text{não depende de } \underset{\sim}{\omega}_1, \ldots, \underset{\sim}{\omega}_k} - \mathbb{E}\underbrace{\left(c_1^2 + \ldots + c_k^2\right)}_{\text{onde } c_i = \underset{\sim}{\omega}_i' \cdot \underset{\sim}{X}}$$

Minimizar $\mathbb{E}\left(\|\underset{\sim}{X}-\underset{\sim}{X}^*\|^2\right)$ é equivalente (36)

a maximizar $\mathbb{E}\left(c_1^2+\dots+c_k^2\right)$ onde $c_i=\underset{\sim}{\omega}_i'\underset{\sim}{X}$.

Temos

$$\mathbb{E}\left(c_1^2+\dots+c_k^2\right)=\mathbb{E}\left(c_1^2\right)+\dots+\mathbb{E}\left(c_k^2\right)$$

e

$$\mathbb{E}\left(c_i^2\right)=\mathbb{E}\left(\left(\underset{\sim}{\omega}_i'\cdot\underset{\sim}{X}\right)^2\right)=\mathbb{E}\left(\left(\underset{\sim}{\omega}_i'\underset{\sim}{X}\right)\left(\underset{\sim}{\omega}_i'\underset{\sim}{X}\right)\right)$$

$$=\mathbb{E}\left(\omega_i'\underset{\sim}{X}\underset{\sim}{X}'\underset{\sim}{\omega}_i\right)=\omega_i'\,\mathbb{E}\left(\underset{\sim}{X}\underset{\sim}{X}'\right)\omega_i$$

$$=\underset{\sim}{\omega}_i\,\Sigma\,\underset{\sim}{\omega}_i$$

Em resumo, precisamos encontrar $\underset{\sim}{w}_1, \underset{\sim}{w}_2, \ldots, \underset{\sim}{w}_k \in \mathbb{R}^p$, ortogonais entre si e com comprimento unitário tal que

$\underset{\sim}{w}_1' \Sigma \underset{\sim}{w}_1 + \underset{\sim}{w}_2 \Sigma \underset{\sim}{w}_2 + \ldots + \underset{\sim}{w}_k \Sigma \underset{\sim}{w}_k$ seja maximizado.

- Vamos considerar apenas o caso $k=2$

$\underset{\sim}{w}_1' \Sigma \underset{\sim}{w}_1 + \underset{\sim}{w}_2 \Sigma \underset{\sim}{w}_2$ ⟵ MAXIMIZAR EM $\underset{\sim}{w}_1$ e $\underset{\sim}{w}_2$ com $\underset{\sim}{w}_1 \perp \underset{\sim}{w}_2$

- Como $\underset{\sim}{w} \Sigma \underset{\sim}{w}$ é máxima quando $\underset{\sim}{w} = \underset{\sim}{v}_1 = 1^{\underline{o}}$ autovetor então

$$\underset{\sim}{w}_1' \Sigma \underset{\sim}{w}_1 + \underset{\sim}{w}_2' \Sigma \underset{\sim}{w}_2 \leq \underbrace{\underbrace{\underset{\sim}{v}_1' \Sigma \underset{\sim}{v}_1}_{\underset{\sim}{v}_1' \lambda_1 \underset{\sim}{v}_1}}_{\lambda_1 \|\underset{\sim}{v}_1\|^2 = \lambda_1} + \underset{\sim}{w}_2' \Sigma \underset{\sim}{w}_2$$

Portanto, tomando $\underset{\sim}{w}_1 = \underset{\sim}{v}_1$, devemos maximizar em $\underset{\sim}{w}_2 \perp \underset{\sim}{v}_1$ e com $\|w_2\| = 1$ a expressão

$$\lambda_1 + \underset{\sim}{w}_2' \Sigma \underset{\sim}{w}_2$$

<u>Solução</u>: $\underset{\sim}{w}_2 = v_2 =$ 2º autovetor. Para ver isto, siga o resto da prova a seguir.

Como $\underset{\sim}{w}_2 \perp \underset{\sim}{v}_1$ e tem expressão única na base de autovetores $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_p$ podemos escrever

$$\underset{\sim}{w}_2 = 0\,\underset{\sim}{v}_1 + c_2\,\underset{\sim}{v}_2 + \ldots + c_p\,\underset{\sim}{v}_p$$

$$= \left[\begin{array}{c|c|c|c} \underset{\sim}{v}_1 & \underset{\sim}{v}_2 & \cdots & \underset{\sim}{v}_p \end{array}\right] \cdot \begin{bmatrix} 0 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$$

$$= P \cdot \begin{pmatrix} 0 \\ c_2 \\ \vdots \\ c_p \end{pmatrix} \quad \text{com } 0^2 + c_2^2 + \ldots + c_p^2 = 1 \text{ pois } \|\underset{\sim}{w}_2\|^2 = 1.$$

Portanto, temos

$$\underset{\sim}{w}_2' \Sigma \underset{\sim}{w}_2 = (0 \ c_2 \cdots c_p) P' \Sigma P \cdot \begin{pmatrix} 0 \\ c_2 \\ \vdots \\ c_p \end{pmatrix}$$

TEOREMA ESPECTRAL

$$= (0 \ c_2 \cdots c_p) P' (P \Lambda P') P \begin{pmatrix} 0 \\ c_2 \\ \vdots \\ c_p \end{pmatrix}$$

$$= (0, \ c_2 \cdots, c_p) \underbrace{P'P}_{=I} \Lambda \underbrace{P'P}_{=I} \begin{pmatrix} 0 \\ c_2 \\ \vdots \\ c_p \end{pmatrix}$$

$$= (0, c_2, \ldots, c_p) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \begin{pmatrix} 0 \\ c_2 \\ \vdots \\ c_p \end{pmatrix}$$

$$= 0 + c_2^2 \lambda_2 + \cdots + c_p^2 \lambda_p \leq c_2^2 \lambda_2 + c_3^2 \lambda_2 + \cdots + c_p \lambda_2$$

$$= \lambda_2 (c_2^2 + \cdots + c_p) = \lambda_2$$

Assim, $\underset{\sim}{w}_2' \Sigma \underset{\sim}{w}_2$ é maximizado com

$\underset{\sim}{w}_2 \perp \underset{\sim}{v}_1 = 1^{\underline{o}}$ ~~au~~autovetor se

$$\underset{\sim}{w}_2' \Sigma \underset{\sim}{w}_2 = \lambda_2 = 2^{\underline{o}} \text{ autovalor.}$$

tomando $\underset{\sim}{w}_2 = \underset{\sim}{v}_2 = 2^{\underline{o}}$ autovetor temos

$$\underset{\sim}{v}_2' \Sigma \underset{\sim}{v}_2 = \underset{\sim}{v}_2' \left(\lambda_2 \underset{\sim}{v}_2\right) = \lambda_2 \underset{\sim}{v}_2' \underset{\sim}{v}_2$$

$$= \lambda_2 \|\underset{\sim}{v}_2\|^2 = \lambda_2$$

Assim, a solução é $\underset{\sim}{w}_2 = \underset{\sim}{v}_2$.

~~QUESTÕES~~ PRA

RESUMO

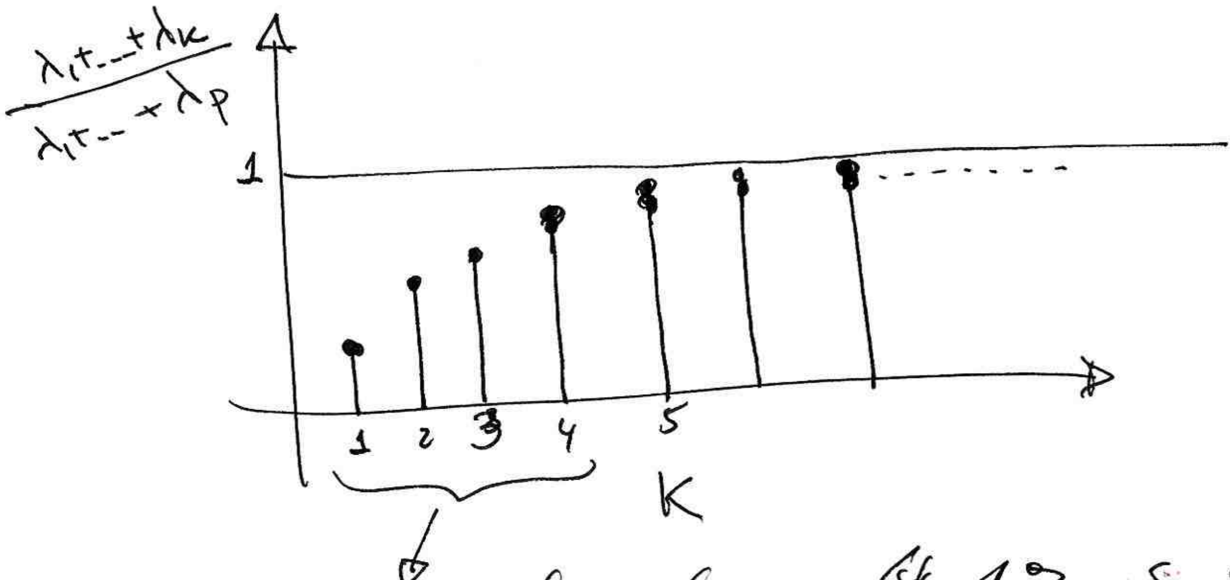
UM VETOR $\underset{\sim}{X} \in \mathbb{R}^p$ PODE SER REDUZIDO DA SEGUINTE FORMA:

$$\underset{\sim}{X} \approx \underset{\sim}{X}^* = c_1 \underset{\sim}{v}_1 + \dots + c_k \underset{\sim}{v}_k$$

onde v_i = i-ésimo autovetor de $\Sigma = \mathbb{V}(\underset{\sim}{X})$

e $\underbrace{c_i = \underset{\sim}{v}_i' \cdot \underset{\sim}{X}}_{\text{i-ÉSIMO COMPONENTE PRINCIPAL}}$

- QUEREMOS $k \ll p$.
- Como escolher k?
- CRITÉRIO prático: plote $\dfrac{\lambda_1 + \dots + \lambda_k}{\lambda_1 + \dots + \lambda_k + \lambda_{k+1} + \dots + \lambda_p}$ versus k. A função é crescente em direção a 1. Quando ficar "achatada", é hora de parar.



possivelmente os 4 1ºs suficientes

Definição

$$\mathbb{V}(\underset{\sim}{X}) = \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_p^2 \end{pmatrix}$$

<u>Variância total</u> $= \sigma_1^2 + \dots + \sigma_p^2 = \mathrm{tr}(\Sigma) = \text{traço}(\Sigma)$

Resultado Variância total $= \sigma_1^2 + \dots + \sigma_p^2 = \lambda_1 + \dots + \lambda_p$

onde λ_i = autovalor de Σ.

Prova: $tr(AB) = tr(BA)$ q.do A e B são quadrados.

Portanto,

pelo teor. espectral

$$\sigma_1^2 + \dots + \sigma_p^2 = tr(\Sigma) = tr(P \Lambda P')$$

$$= tr(\underbrace{P'P}_{=I}\Lambda) = tr(\Lambda) = tr\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix}$$

$$= \lambda_1 + \dots + \lambda_p$$

$Y_i = \underset{\sim}{v}_i' \underset{\sim}{X}$ = i-ésimo componente principal

$$\mathbb{V}(Y_i) = \mathrm{Var}(\underset{\sim}{v}_i' \underset{\sim}{X}) = \underset{\sim}{v}_i' \underbrace{\Sigma \underset{\sim}{v}_i}_{\lambda_i \underset{\sim}{v}_i} = \lambda_i \underset{\sim}{v}_i' \underset{\sim}{v}_i = \lambda_i \|v_i\|^2$$
$$= \lambda_i$$

Assim

Variância total $= \sigma_1^2 + \ldots + \sigma_p^2 = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_p)$

$= \lambda_1 + \ldots + \lambda_p = \mathrm{Var}(Y_1) + \ldots + \mathrm{Var}(Y_p)$

= Soma das variâncias dos Componentes principais

- Proporç da variância total devida ao i-ésimo componente principal

$$= \frac{\lambda_i}{\lambda_1 + \dots + \lambda_p} = \frac{\lambda_i}{\sigma_1^2 + \dots + \sigma_p^2} = \frac{\lambda_i}{\left(\text{Variância total}\right)}$$

- $$\frac{\lambda_1 + \dots + \lambda_k}{\lambda_1 + \dots + \lambda_k + \dots + \lambda_p} = \frac{\lambda_1 + \dots + \lambda_k}{\left(\text{variância total}\right)} =$$

= proporç da variância total "explicada" pelos K 1ºs componentes principais

(47)

- Os componentes principais são afetados por mudanças de escala.
- Trabalhar com dólares ou reais implicam em Principal Components diferentes.
- trabalhar com graus Celsius ou ~~Fahr~~ ~~Fah~~ Fahrenheit ?
- Quando as escalas dos variáveis são bem diferentes ou se elas possuem uma faixa de variação muito diferente, é comum trabalharmos com as variáveis padronizadas

Isto é, trabalhamos com o vetor $\underset{\sim}{Z}$ onde

$$\underset{\sim}{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{pmatrix} = \begin{pmatrix} \frac{X_1 - \mu_1}{\sigma_1} \\ \frac{X_2 - \mu_2}{\sigma_2} \\ \vdots \\ \frac{X_p - \mu_p}{\sigma_p} \end{pmatrix}$$

Precisamos da matriz de <u>Covariância</u> do vetor $\underset{\sim}{Z}$

$$Cov\left(\underset{\sim}{Z}\right) = \begin{bmatrix} Var(Z_1) & Cov(Z_1, Z_2) & \cdots & Cov(Z_1 Z_p) \\ Cov(Z_2 Z_1) & Var(Z_2) & \cdots & Cov(Z_2, Z_p) \\ \vdots & & & \\ Cov(Z_p Z_1) & Cov(Z_p Z_2) & \cdots & Var(Z_p) \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & & \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix} =$$

= matriz de correlação de $\underset{\sim}{X}$

PCA NA PRÁTICA

$\boxed{\text{Tabela de dados}}$

$n \times p$

n itens

p variáveis

$$D = \begin{bmatrix} y_1 \\ \hline y_2 \\ \hline \vdots \\ \hline y_n \end{bmatrix} = \begin{bmatrix} g_1 & \vdots & g_2 & \vdots & \cdots & \vdots & g_p \end{bmatrix}$$

n linhas

p colunas

⊕ Estime a matriz de covariância $\underset{p \times p}{\Sigma}$ dos p variáveis

⊕ $S = \text{cov}(D)$ em R