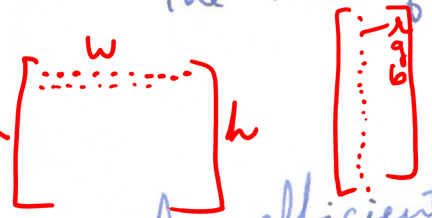
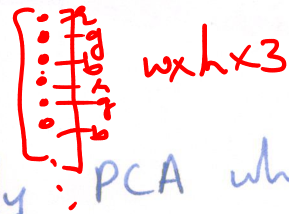
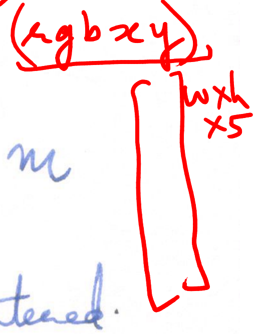
The proof shows that setting the columns of matrix $\underline{U}$ as the n leading eigenvectors of A will ensure that

$$\text{trace}(\underline{U}^T A \underline{U}) = \sum_{i=1}^{n} \lambda_i$$

It can be showed that the reconstruction loss is the sum of discarded eigenvalues $\sum_{i=n+1}^{d} \lambda_i$. distortion







An efficient way to apply PCA when $d \gg m$

Consider vectors $\underline{x}_i$ which have been mean-centered.

The covariance matrix $\underline{C} = \underline{X}^T \underline{X}$ where $\underset{m \times d}{\underline{X}} =$ data matrix.

Clearly, $\underline{C}_{d \times d}$ is very large. 10^6

Instead, we formulate another matrix $\underset{m \times m}{\underline{B}} = \underline{X}\,\underline{X}^T$ $\quad m{\times}d \times d{\times}m$

$\underline{B}$ is $m \times m$ matrix.

The $(i,j)^{th}$ element of B is given as $\boxed{\langle \underline{x}_i, \underline{x}_j \rangle}$

$\underline{B}$ is also called as the Gram matrix.

We obtain the eigen decomposition of $\underline{B}$. $\boxed{\underline{B}\,\underline{u} = \lambda \underline{u}}$

If $\underline{u}$ is an eigenvector of $\underline{B}$ then

$$\underbrace{(\underline{X}\,\underline{X}^T)}_{B}\underline{u} = \lambda \underline{u}$$

$$X^T(\underline{X}\,\underline{X}^T)\,\underline{u} = \lambda(\underline{X}^T \underline{u})$$

since $C = \underline{X}^T \underline{X}$, we have

$\underline{C\nu = \lambda\nu}$

$$C(X^T u) = \lambda(X^T u).$$

Thus, $\left(\frac{X^T u}{\|X^T u\|}\right)$ is a normalized eigenvector of ~~A~~ C with eigen value λ.

Thus we have calculated the PCA solution for C using the eigendecomposition of matrix B.

$m=100$ images $w \times h$.

$d = w \times h \times 3$

$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{100} \end{bmatrix}_{100}$ $m=100$ eigen vectors

Eigen vectors in a d-dim space.

How many eigen vectors will we get ??

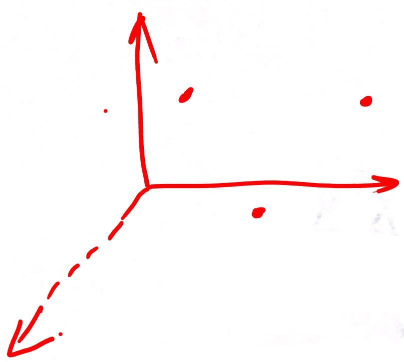
with non-zero eigen values

3-d space

2 points in a 3d space.

m points in a d-dim space

$m \ll d$



3 points in a 3 d space

plane

2 eigen vectors.

3×10^6 to 100

bases

image

$$I_1 = \alpha_1 \underbrace{\begin{bmatrix} - \\ \vdots \\ - \\ \vdots \end{bmatrix}}_{E_1\ d\times1} + \alpha_2 \underbrace{\begin{bmatrix} - \\ - \\ - \\ \vdots \end{bmatrix}}_{E_2\ d\times1} + \cdots + \alpha_{100} \underbrace{\begin{bmatrix} - \\ - \\ - \\ \vdots \end{bmatrix}}_{E_{100}\ d\times1}$$

$E_{m'}$

$d = w \times h \times 3$

I_1 in the new space eigenspace with 100 dimensional vector $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{100} \end{bmatrix}$

$I_1 \ldots I_{100}$ each image requires a 100-dim vector.

$\underbrace{3\times10^6}_{d} \rightarrow d = 100$

We need to store the bases of the new space also

$m = 100$

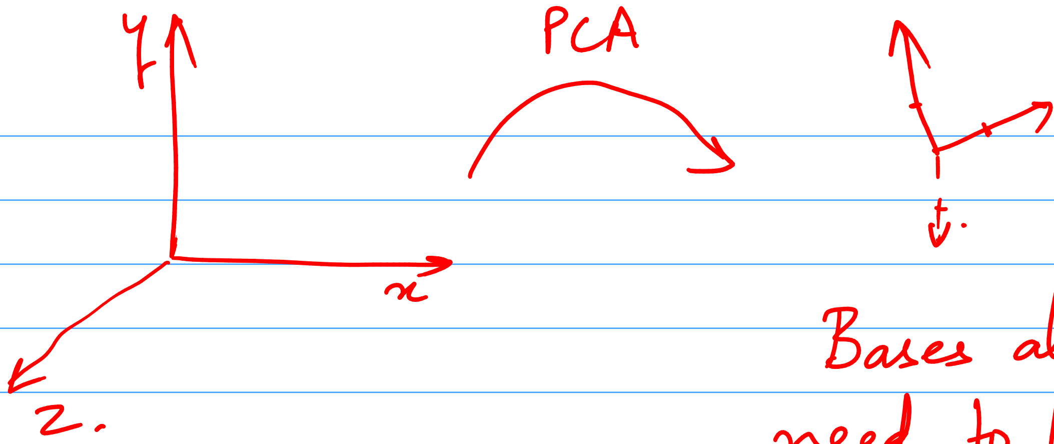
$$\underbrace{\underbrace{m\times100}_{} + m\times w\times h\times3}_{\text{Eigen space}} \quad \Bigg| \quad \underbrace{m\times w\times h\times3}_{\text{original.}} \text{ color values}$$

Using $m' < m$ eigenvectors in the new space

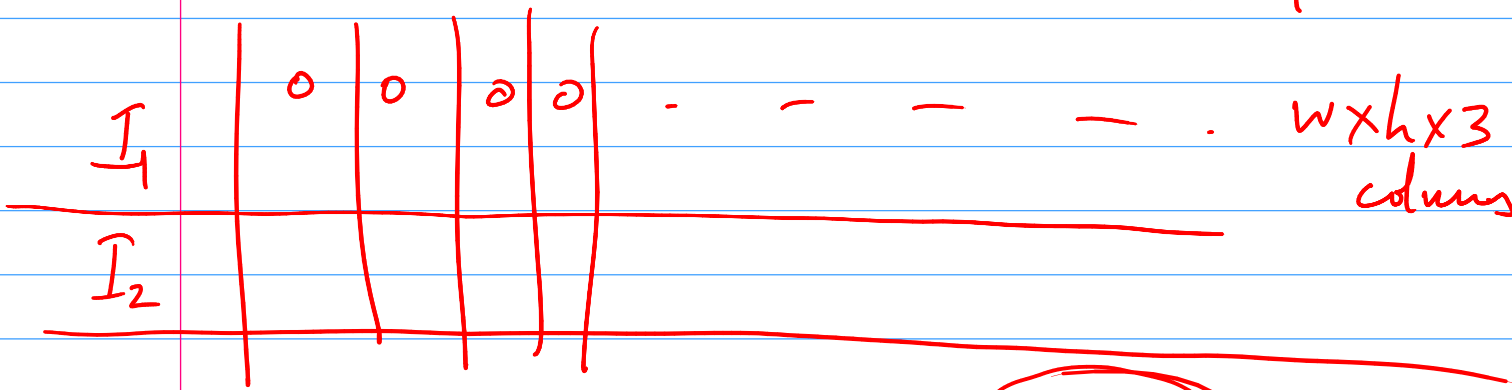
$m' = 50$

$m\times50 + 50\times w\times h\times3$

NO Image is still $w\times h\times3$.

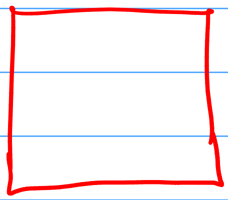


Bases also need to be stored.



$w \times h \times 3$ columns

CSV



Tensors $w \times h \times 3$ channels