PCA and the Singular Value Decomposition (SVD)

Hervé Abdi

— PCA and Beyond

www.utdallas.edu/∼herve herve@utdallas.edu

1. The Real Tool: The SVD

- SVD is for Singular *V*alue *D*ecomposition
- Generalizes the eigen-decomposition
- 1. The Geometry of the SVD (with the words example)

  . . . and an illustration with images
- 2. The big questionWhat does the SVD optimizes?

2. Images of the Components Building Back!

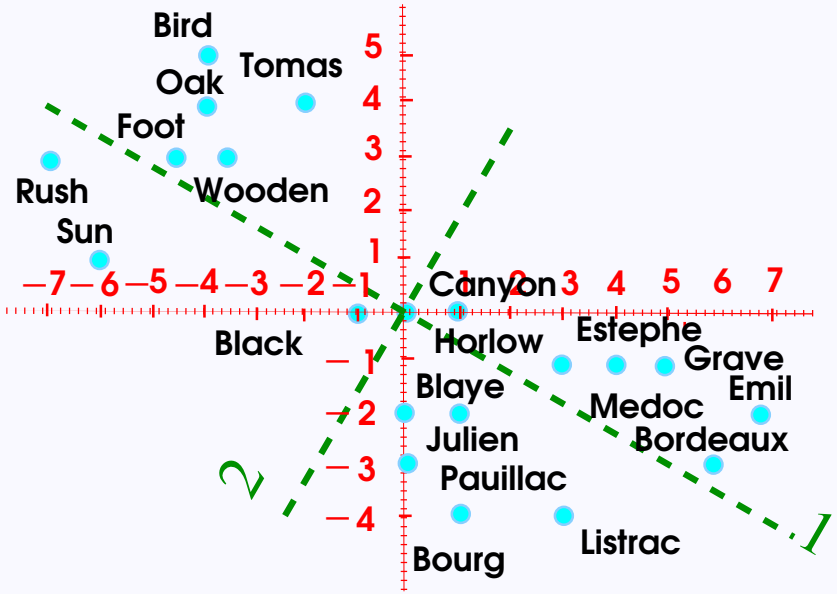
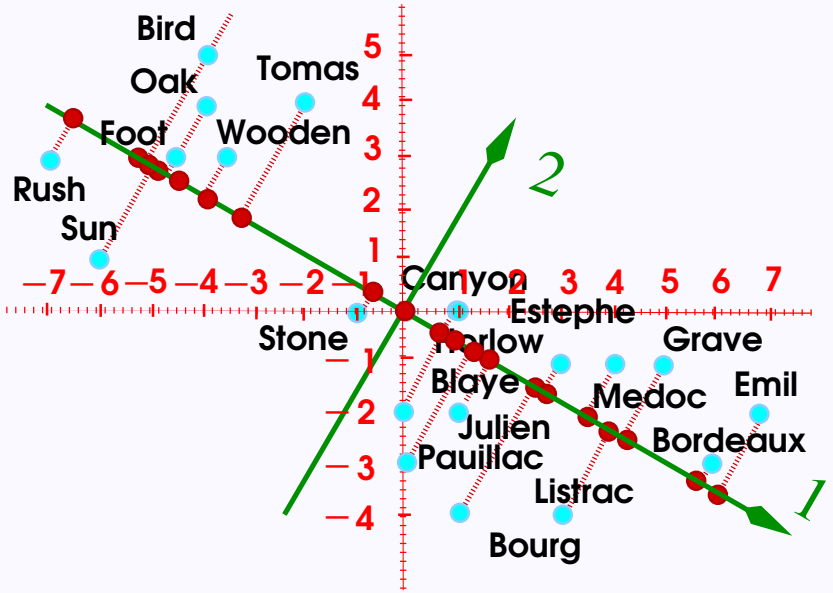
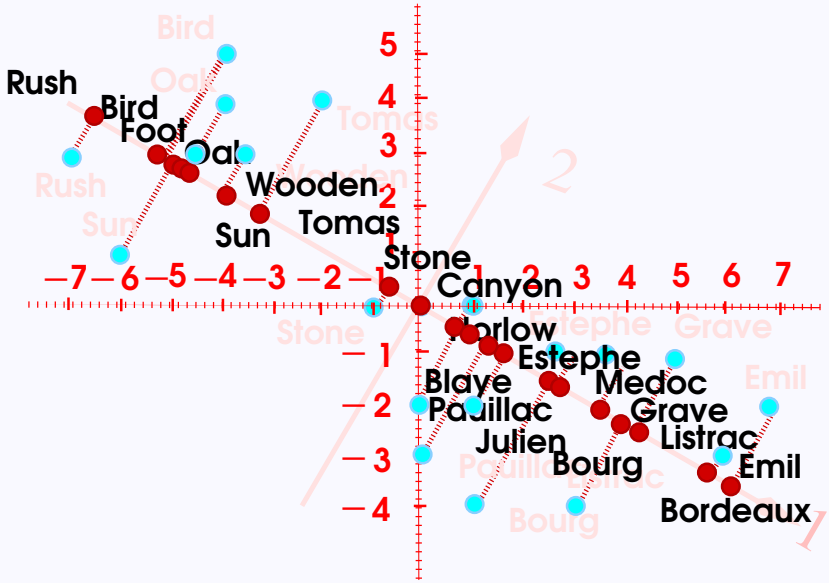
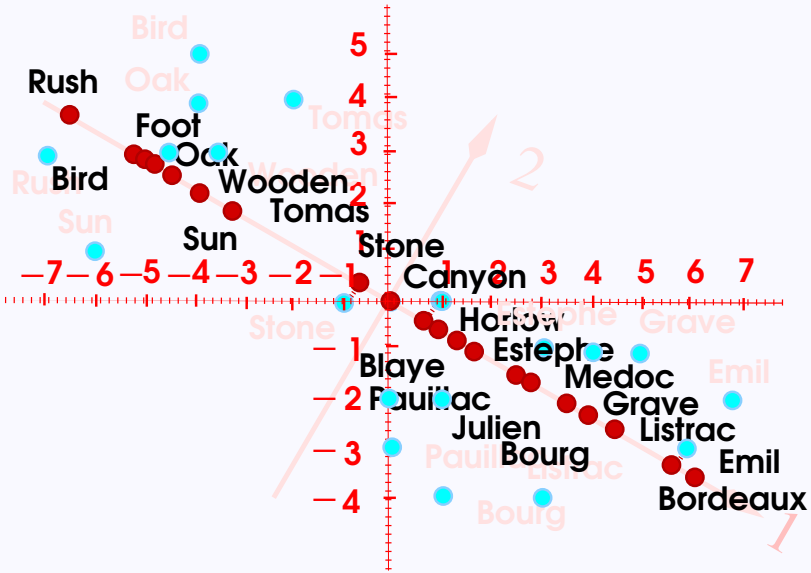


Figure 1: The Components in the space of the original variables.

The data according to the first component







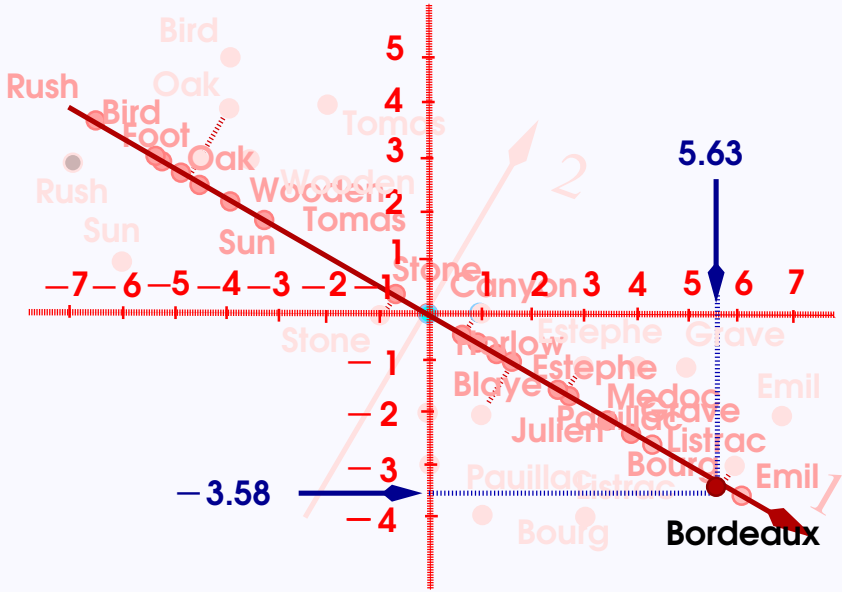
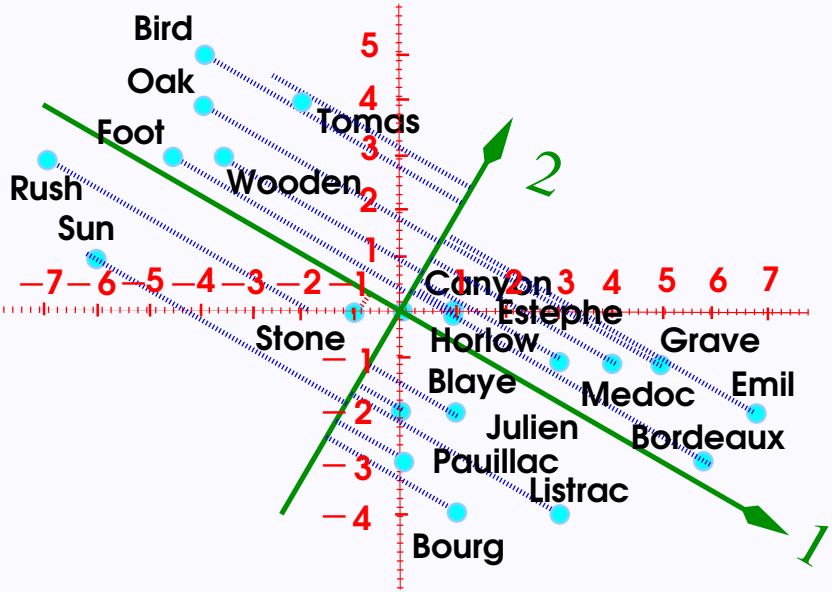
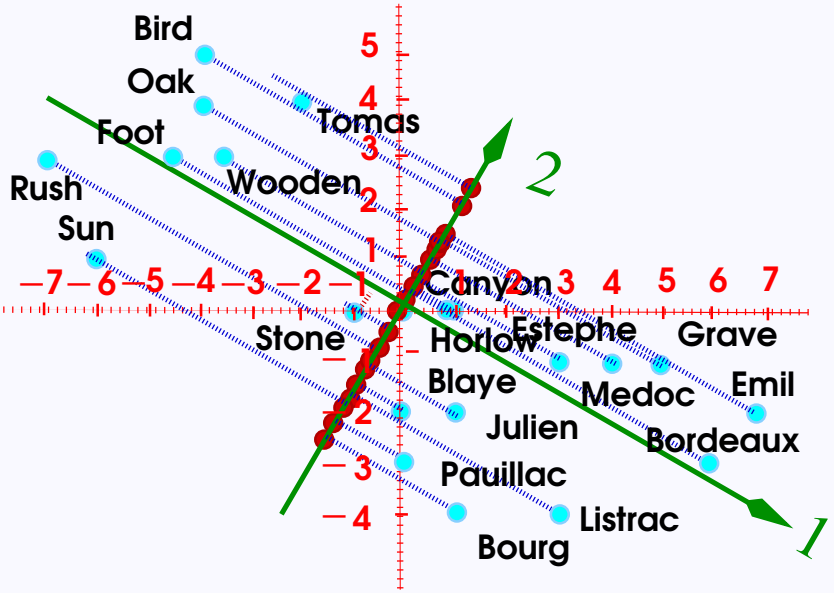
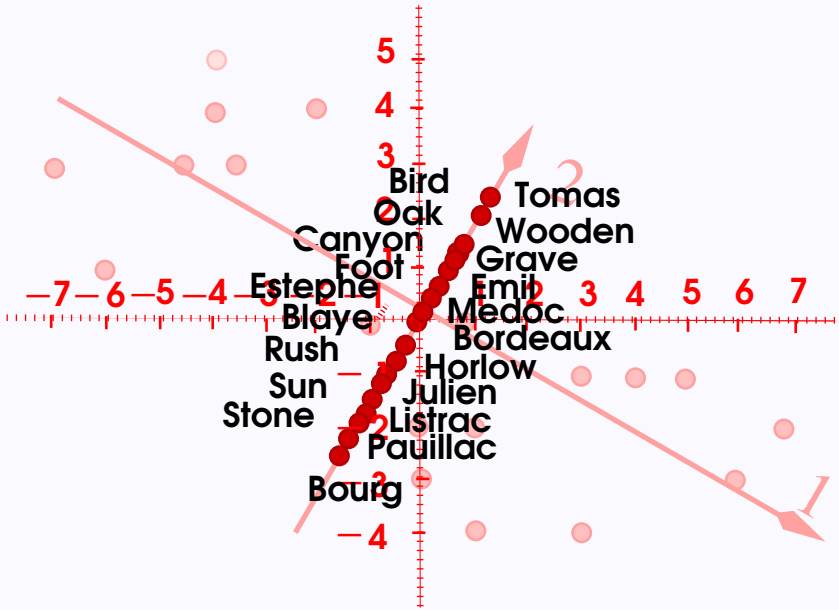


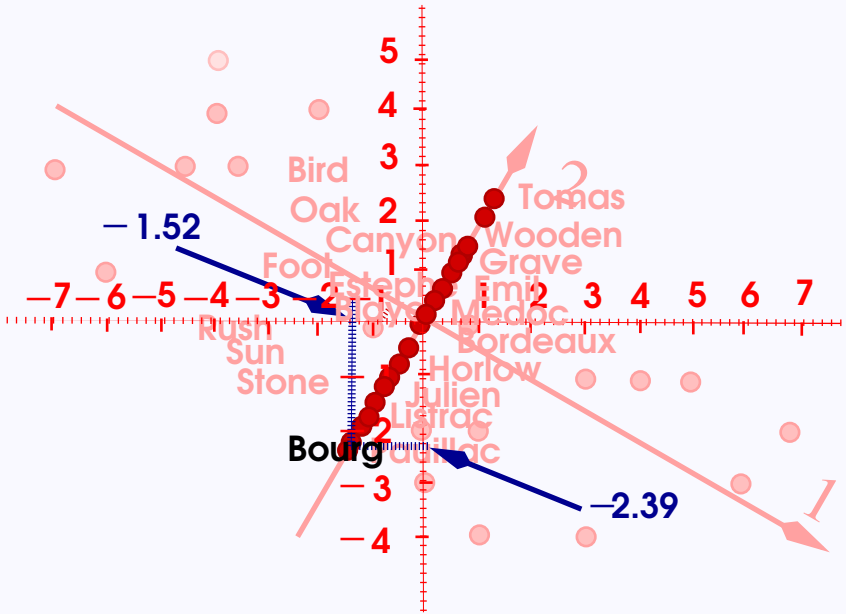
Table 1: Building back the data table one component at a time. Original variables, F_1, data table according to F_1 F_2, data table according to F_2, $\widehat{w}_1$ is w according to F_1. The original data table is the sum of the "simple data tables."

	w	y	F_1	$\widehat{w}_1$	$\widehat{y}_1$	F_2	$\widehat{w}_2$	$\widehat{y}_2$
Bordeaux	6	−3	6.67	5.63	−3.58	0.69	0.37	0.58
Black Stone	−1	0	−0.84	−0.71	0.45	−0.54	−0.29	−0.45
Listrac	3	−4	4.68	3.95	−2.51	−1.76	−0.95	−1.49
Canyon Creek	1	0	0.84	0.71	−0.45	0.54	0.29	0.45
Côte de Bourg	1	−4	2.99	2.52	−1.61	−2.84	−1.52	−2.39
Foot Hill	−4	3	−4.99	−4.21	2.68	0.38	0.21	0.32
Horlow	0	0	0.00	0.00	0.00	0.00	0.00	0.00
St Estephe	3	−1	3.07	2.59	−1.65	0.77	0.41	0.65
Wooden Hill	−3	3	−4.14	−3.49	2.22	0.92	0.49	0.78
Côte de Blaye	0	−2	1.07	0.91	−0.58	−1.69	−0.91	−1.42
Sun Set	−6	1	−5.60	−4.72	3.01	−2.38	−1.28	−2.01
Black Bird	−4	5	−6.06	−5.11	3.25	2.07	1.11	1.75
Médoc	4	−1	3.91	3.30	−2.10	1.30	0.70	1.10
St Julien	1	−2	1.92	1.62	−1.03	−1.15	−0.62	−0.97
Pauillac	0	−3	1.61	1.36	−0.86	−2.53	−1.36	−2.14
Gold Rush	−7	3	−7.52	−6.34	4.04	−1.23	−0.66	−1.04
Oak Vill	−4	4	−5.52	−4.66	2.96	1.23	0.66	1.04
Grave	5	−1	4.76	4.01	−2.55	1.84	0.99	1.55
St Emilion	7	−2	6.98	5.89	−3.75	2.07	1.11	1.75
Tomasello	−2	4	−3.83	−3.24	2.06	2.30	1.24	1.94

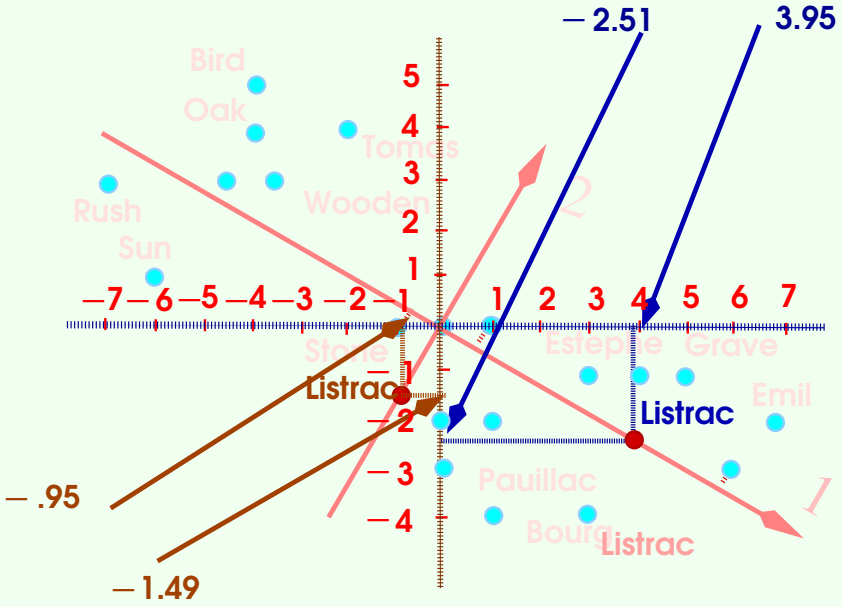


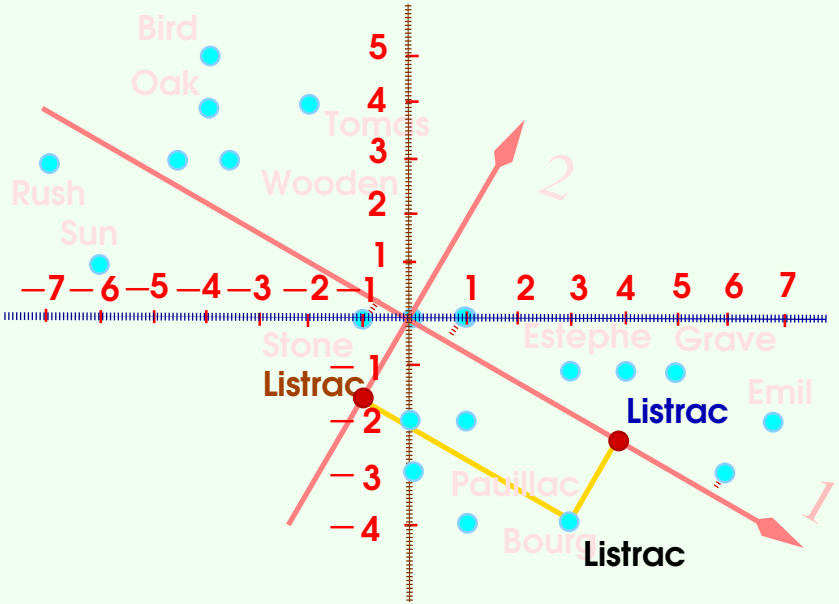


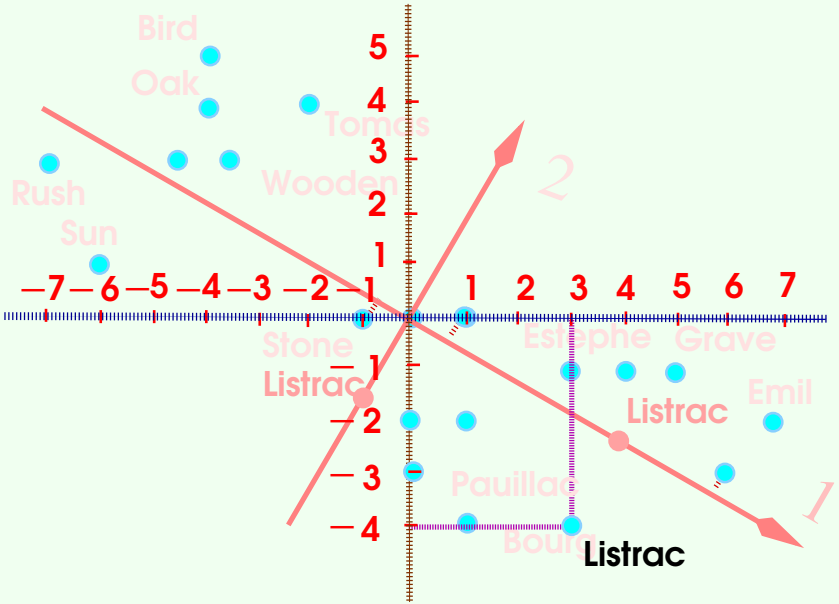




3. Building Back the Data: Addition







4. The return of Cosines

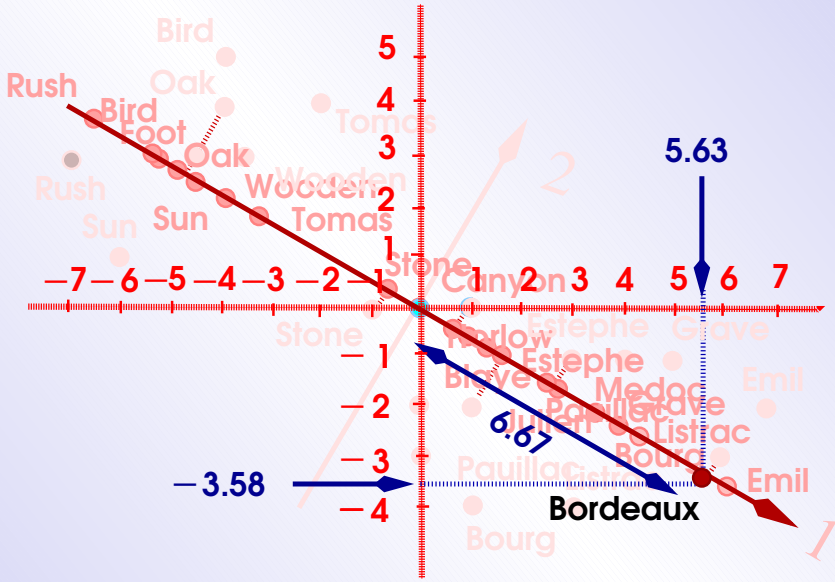


Figure 2: The coefficient of proportionality used to convert the value of F_1 into $\widehat{w}_1$ or $\widehat{y}_1$ is the cosine of the angle between the component and the original variables.

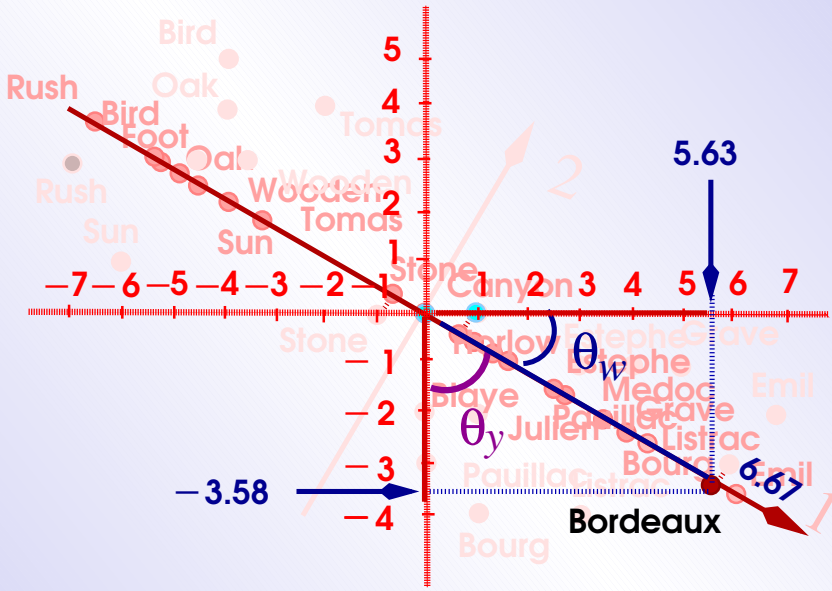


Figure 3: The coefficient of proportionality used to convert the value of F_1 into $\widehat{w}_1$ or $\widehat{y}_1$ is the cosine of the angle between the component and the original variables.

5. The Singular Value Decomposition: Take One!

$$\widehat{w}_{\text{bordeaux},1} = .8437 \times F_{\text{bordeaux},1} = .8437 \times (6.6726) = 5.6294 \, .$$

$$\widehat{y}_{\text{bordeaux},1} = -.5369 \times F_{\text{bordeaux},1} = -.5369 \times (6.6726) = -3.5824$$

these are also cosines:

$\vartheta_{w,1} = 32^o$ (*i.e.*, $\cos(32) = .8437$)

$\vartheta_{y,1} = 122^o$ (*i.e.*, $\cos(122) = .5369$)

$$\widehat{w}_{\text{bordeaux},2} = .5369 \times F_{\text{bordeaux},2} = .5369 \times (0.6903) = 0.3706 \, .$$

$$\widehat{y}_{\text{bordeaux},1} = .8437 \times F_{\text{bordeaux},2} = .8437 \times (0.6903) = 0.5824 \, .$$

6. Cosine: the weights to build back everything

6.1. F from w and y

$w_{\text{bordeaux}} = 6$ and $y_{\text{bordeaux}} = -3$ as

$$F_{\text{bordeaux},1} = .8437 \times 6 + (-.5369) \times (-3) = 5.0620 + 1.6106 = 6.6726.$$

6.2. w and y from F

$F_{\text{bordeaux},1} = 6.6729$ and $F_{\text{bordeaux},2} = 0.6903$ as

$$w_{\text{bordeaux}} = .8437 \times 6.6726 + .5369 \times 0.6903 = 5.6294 + 0.3706 = 6.$$

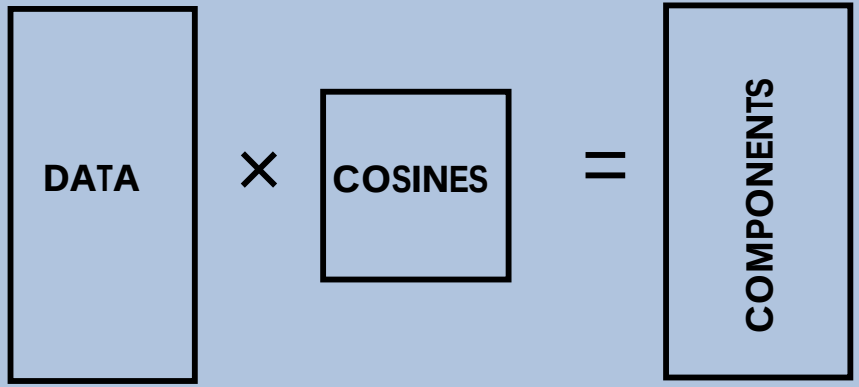


Figure 4: **F** from **X** and **Q**

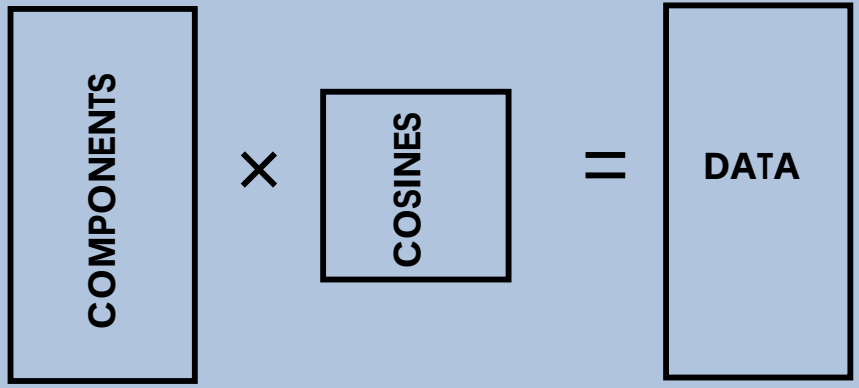


Figure 5: X from F and Q

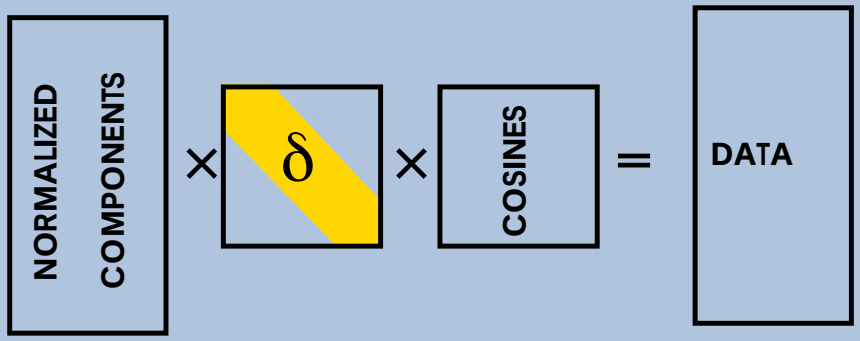


Figure 6: The SVD! $\mathbf{P\Delta Q}^T = \mathbf{X}$

7. What about these cosines and the loadings?

$$\text{Cosines} = \text{Loadings} \times \sqrt{\frac{SS_w}{\lambda}}$$

8. Face example: Take 1

Figure 7: The SVD! A 204 rows $\times$ 290 columns = $59{,}610$ pixels image.

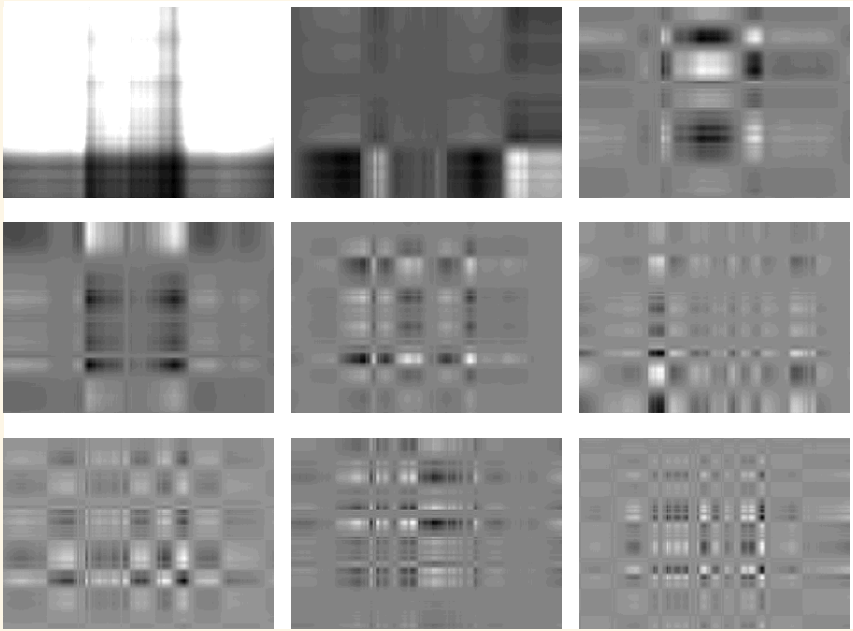


Figure 8: The SVD!

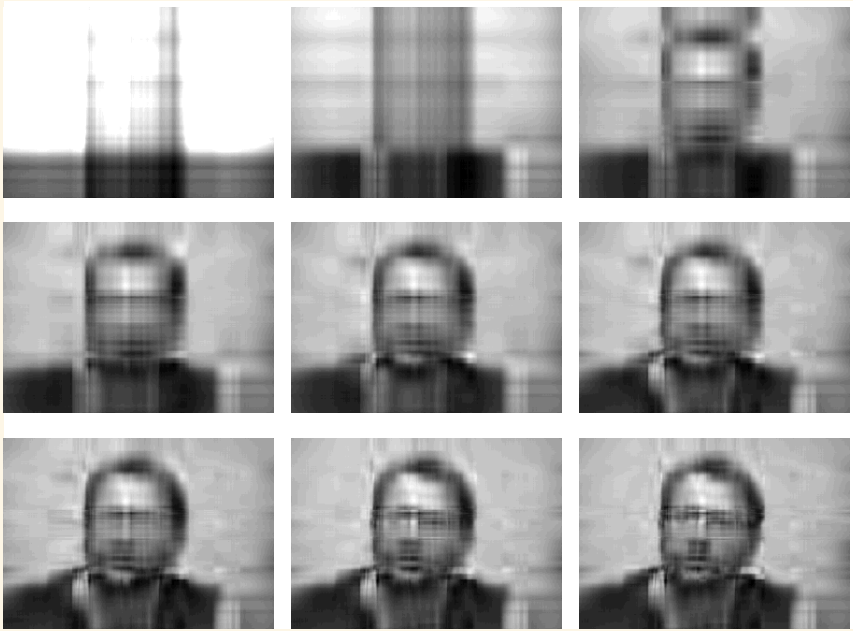


Figure 9: The SVD!

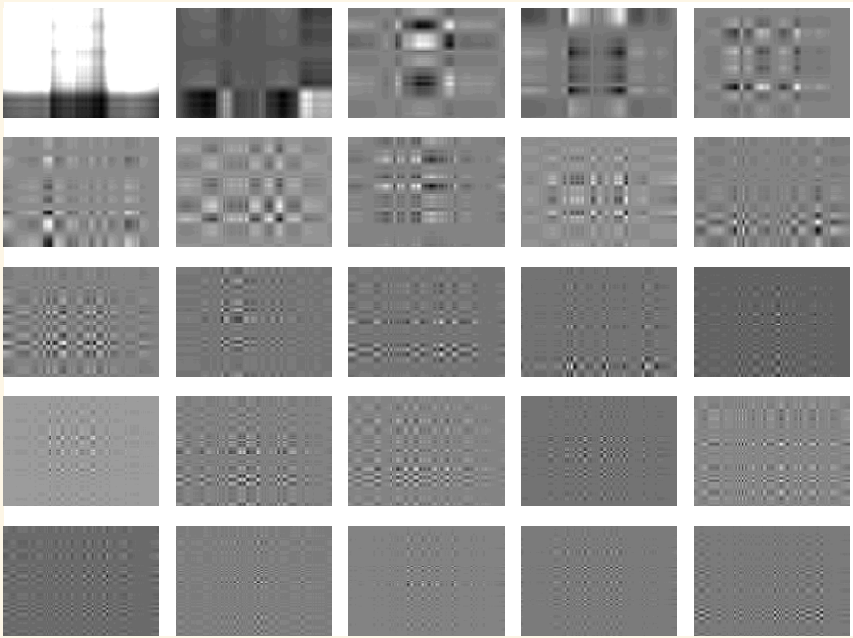


Figure 10: The SVD!

Figure 11: The SVD!

Figure 12: The SVD! Compression rate is around 80%.

9. So what does the SVD maximize?

Remember: $\mathbf{X} = \mathbf{P}\mathbf{\Delta}\mathbf{Q}^T$

So we want to find 2 vectors $\mathbf{p}$ and $\mathbf{q}$ such that

$\widehat{\mathbf{X}} = \delta \mathbf{p}\mathbf{q}^T = \mathbf{f}\mathbf{q}^T$ with $\mathbf{p}^T\mathbf{p} = \mathbf{q}^T\mathbf{q} = 1$ and $\mathbf{f} = \delta\mathbf{p}$

(i.e., $\widehat{\mathbf{X}}$ is a rank one matrix) with

- $\arg\max_{\mathbf{p},\mathbf{q}} = \delta = \text{trace}\left(\mathbf{X} \times \widehat{\mathbf{X}}^T\right)$
- under the constraints that: $\mathbf{p}^T\mathbf{p} = \mathbf{q}^T\mathbf{q} = 1$

These are δ_1, $\mathbf{p}_1$, and $\mathbf{q}_1$.

- Now "deflate" $\widehat{\mathbf{X}}$
- Get $\mathbf{X}^{[1]} = \mathbf{X} - \delta_1 \mathbf{p}_1 q_1^T$
- Now get δ_2, $\mathbf{p}_2$, and $\mathbf{q}_2$ from $\mathbf{X}^{[1]}$ . . .
- Keep on doing till there is nothing left of $\mathbf{X}$
- Et voilà!!