Project 2 - Appendix 1

Amélia O. F. da S.
190037971@unb.br

Departamento de Ciência da Comptutação
Universidade de Brasília
Campus Darcy Ribeiro, Asa Norte
Brasília-DF, CEP 70910-900, Brazil

1 Describing the maximisation problem

1.1 Determinants and scatter

The determinant of a scatter matrix can be interpreted as a measurement of the total "dispersion" of that set of data[1].

1.2 LDA

Given the scatter matrix for variation "between" classes S_b and the scatter matrix for variation "within" classes S_w and assuming the derivative represents the scatter, the ratio to be maximised can be derived as follows:

$$\begin{aligned} r &= \frac{|S_b|}{|S_w|} \\ r &= |S_b||S_w|^{-1} = |S_b||S_w^{-1}| \\ r &= |S_b\ S_w^{-1}| \end{aligned} \tag{1}$$

The problem of maximising a matrix's determinant using a linear transformation WAW^T is equivalent to finding the roots of its characteristic polynomial (finding its eigenvalues)[2]. Therefore, maximising the $\frac{|S_b|}{|S_w|}$ ratio through a linear transformation is equivalent to finding the *Eigenvectors* W of $S_b\ S_w^{-1}$.

1.3 PCA

Following the same logic, maximising the scatter $|S|$ with a linear transformation is equivalent to finding the *Eigenvectors* W of S.

2 Example situation for nearest-mean classification misclassifications

Using LDA to improve nearest-mean classifications might seem redundant at first thought, but there are situations where it fails when used in the original data basis vectors. This generally occurs when variation within axes is significantly different for each of one.

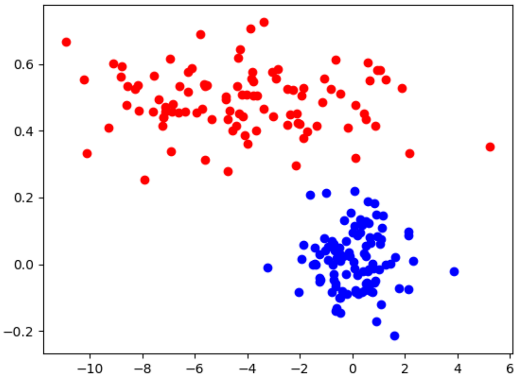

Figure 1: Heavily asymmetrical variation distribution across axes

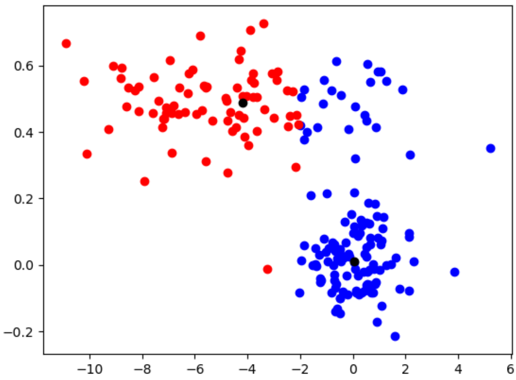

Figure 2: Nearest mean misclassifications on the original projection

3 Misclassifications and projections

Using single-axis projections and thresholding yielded 4 total misclassifications. Analysing the final axis projections allows us to see where the linear classifier failed.

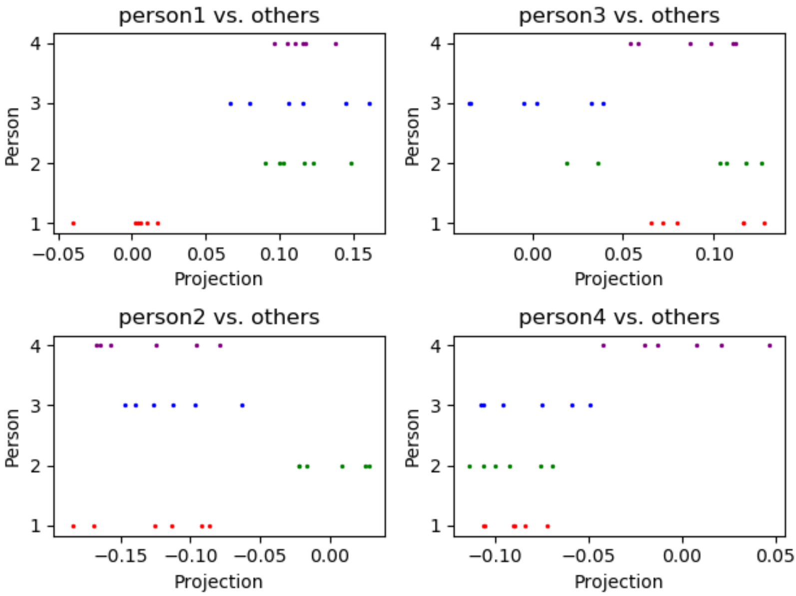

Figure 3: One-against-all axis projections for the four subjects

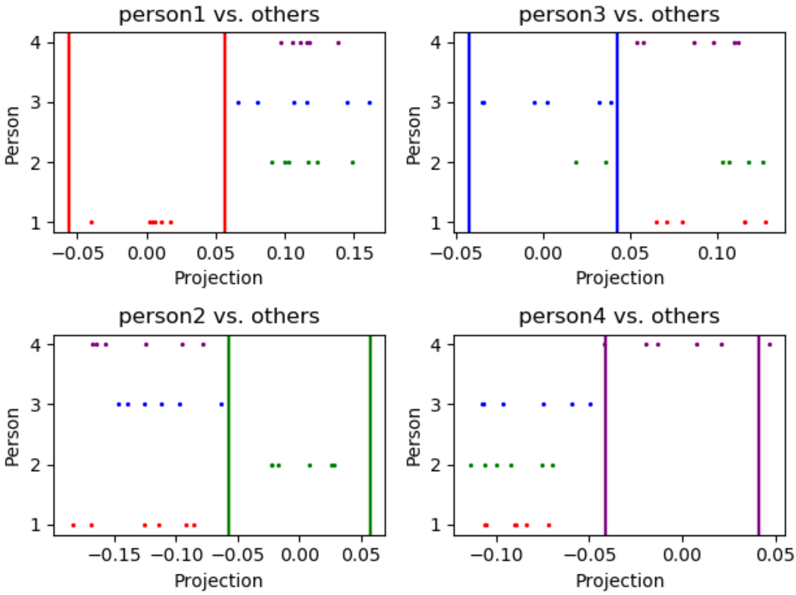

Figure 4: One-against-all axis projections for the four subjects (including threshold visualisations)

Person 3's projection includes two samples from Person 2's set. Person 4's projection excludes 2 of Person 4's own samples.

Improving the threshold calculation and using a more robust classifier would likely solve those problems.

4 Visual differences evidenced by projection weights

By comparing each person's most important features with sample projections it's possible to visually understand why those characteristics were chosen for the resulting discrimination.

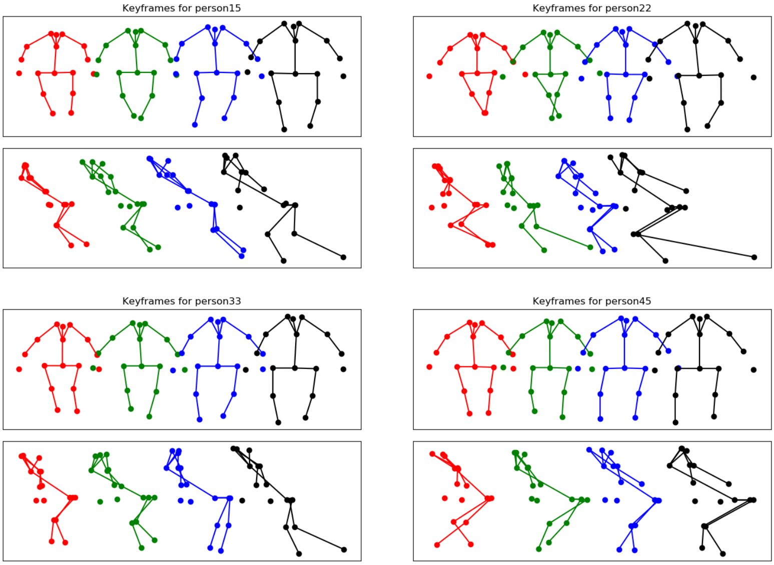

Figure 5: Samples from each of the four subjects

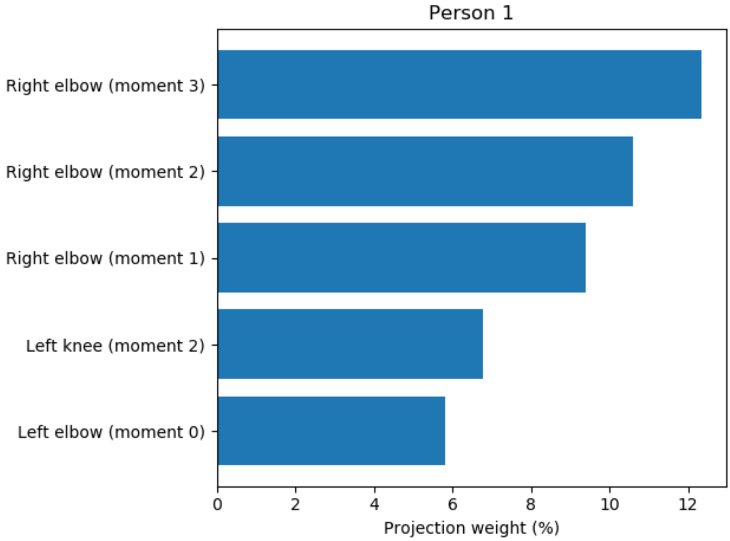

Figure 6: Person 1's feature importances

Person 1's sample is indeed different from the others on these criteria: their right elbow and left knee stay at a shallower angle than the others.

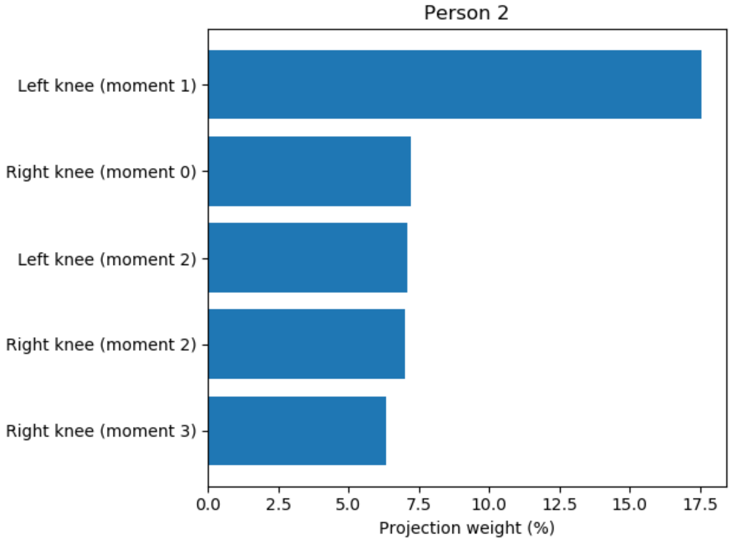

Figure 7: Person 2's feature importances

Person 2 presents a very idiosyncratic walk pattern, keeping both feet almost in line with each other. This was reflected on their projection axis' feature importances, with knee angles dominating the most important features.

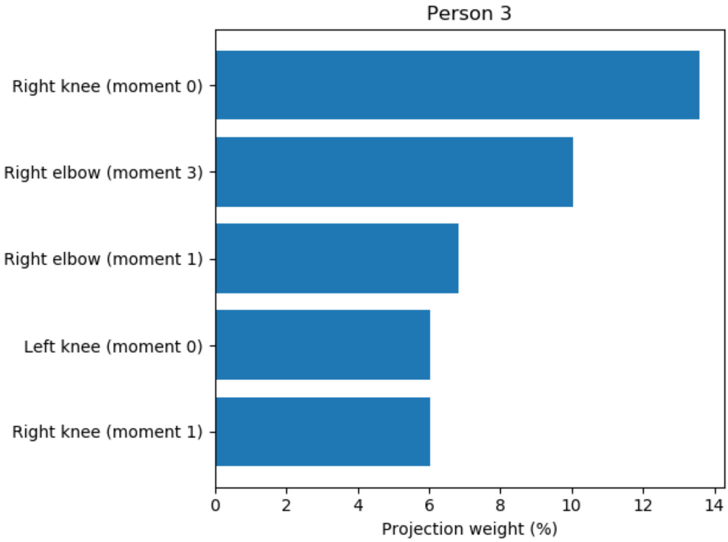

Figure 8: Person 3's feature importances

Person 3's main feature is their left knee's position at moment 0 (the leftmost/red skeleton on the sample images). It's bent at a shallower angle than all other subjects' samples.

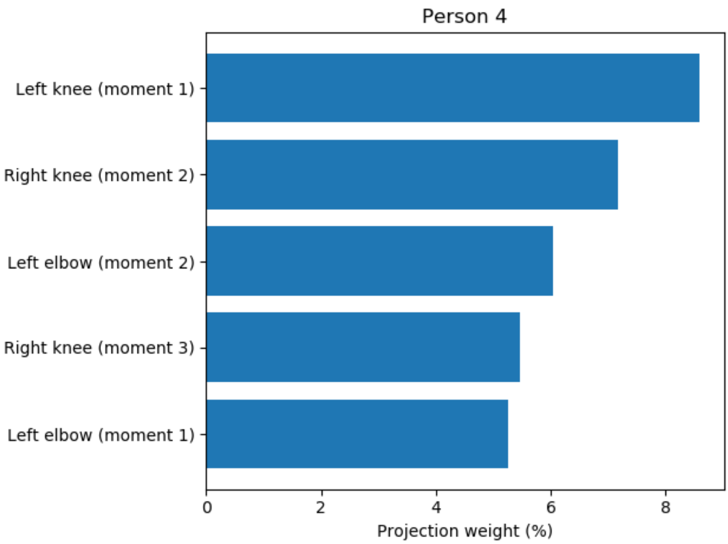

Figure 9: Person 4's feature importances

Person 4 didn't have many peculiar characteristics on their gait, which reflected as a more uniform weight distribution for all features. Their most important feature comprised approximately 8% of the projection weight, whilst for the other projections, the most important feature was always above 12% of the projection weight.

References

[1] W. Härdle and L. Simar. Applied multivariate statistical analysis. 2003.

[2] Lieven Vandenberghe, Stephen Boyd, and Shao-Po Wu. Determinant maximization with linear matrix inequality constraints. *SIAM journal on matrix analysis and applications*, 19(2):499–533, 1998.