Partial and Weighted Jacobi

Jan de Leeuw - University of California Los Angeles

Started November 07, 2023, Version of November 18, 2023

Abstract

TBD

Contents

Note: This is a working paper which will be expanded/updated frequently. All suggestions for improvement are welcome.

1 Introduction

De Leeuw and Pruzansky (1978) De Leeuw and Ferrari (2008) De Leeuw (2017) De Leeuw (2018)

2 Just the formulas

Suppose A is a symmetric matrix of order n and W is a symmetric zero-one matrix of order n. We call A the *target* and W the *pattern*. A pattern is *hollow* if it has zero diagonal and *complete* if all off-diagonal elements are positive.

Using target A and pattern W we define the loss function

$$\sigma(k_1,\cdots,k_n):=\frac{1}{4}\sum_{i=1}^n\sum_{j=1}^n w_{ij}(k_i'Ak_j)^2 \tag{1}$$

Space $\mathcal{K}_{n_1}\oplus\cdots\oplus\mathcal{K}_{n_p}$

Here $(k_1,\cdots,k_n)$ are the n columns of the matrix $K\in\mathcal{K}_n$, the space of all square orthonormal matrices (those with $K'K=KK'=I$, also known as the *rotation* matrices). Thus $k_i'k_j=\delta^{ij}$, where the Kronecker delta δ^{ij} is equal to one if $i=j$ and equal to zero otherwise.

The problem $\mathbb{P}(A,W)$ we study in this report is computing the minimum and a minimizer of σ over $K\in\mathcal{K}_n$. This minimum always exists, because the set $\mathcal{K}_n$ of rotation matrices is compact and the loss function σ is continuous and bounded below by zero. The minimum, and the minimizer, are not necessarily unique.

Since there always exists a K for which $K'AK$ is diagonal it follows that the minimum of σ is equal to zero whenever the pattern is hollow. In that case, any complete orthonormal set of eigenvectors of A is a minimizer of @ref(eq.loss). This result is independent of the off-diagonal elements of W.

The minimization problem $\mathbb{P}(A,W)$ also includes the case in which we minimize over $p<n$ vectors k_i, or equivalently over$K\in\mathcal{K}_{np}$, the Stiefel manifold of all $n\times p$ matrices with $K'K=I$. Simply choose the pattern

$$\begin{bmatrix}W_{p\times p}&0_{p\times(n-p)}\\0_{(n-p)\times p}&0_{p\times(n-p)}\end{bmatrix}, \tag{2}$$

for which

$$\sigma(K)=\frac{1}{4}\sum_{i=1}^p\sum_{j=1}^p w_{ij}(k_i'Ak_j)^2. \tag{3}$$

If $k_1,\cdots,k_p$ is any set of p orthormal eigenvectors with eigenvalues $\lambda_1,\cdots,\lambda_p$ then

$$\sigma(K)=\frac{1}{4}\sum_{i=1}^p w_{ii}\lambda_i^2, \tag{4}$$

which is of course zero for hollow patterns. This there are many minima in this case, all with the same function value zero.

It should be emphasized that most of our results and formulas remain true if W is not binary but non-negative.

3 Derivatives

To get more insight into the loss function (1) we compute its first and second derivatives.

The partials with respect to k_s are

$$\mathcal{D}_s\sigma(k_1,\cdots k_p)=A\sum_{\ell=1}^n w_{s\ell}(k_s'Ak_\ell)k_\ell. \tag{5}$$

If A is non-singular then $\mathcal{D}_s\sigma(k_1,\cdots k_p)$ is zero if and only if $w_{s\ell}(k_s'Ak_\ell)=0$ for all ℓ. If A is non-singular and W is hollow and complete then $\mathcal{D}_s\sigma(k_1,\cdots k_p)=0$ if and only if $k_s'Ak_\ell=0$ for all $\ell\neq s$. Thus Ak_s must be orthogonal to all k_ℓ with $\ell\neq s$, which means that $Ak_s=\lambda_s k_s$, and thus k_s is an eigenvector of A with eigenvalue λ_s.

If A is singular, say of rank $r<n$, then we can use a basis L for the non-null space of A and a basis L_0 for the null space of A. $k_i=Lt_i+L_0s_i$ then $k_i'Ak_j=t_i'L'ALt_j$ and $L'AL$ is non-singular. Thus

$$\sigma(k_1,\cdots,k_n)=\frac{1}{4}\sum_{i=1}^n\sum_{j=1}^n w_{st}(t_i'L'ALt_j)^2$$

which must be minimzied over the $t_i$of length r.

The Hessian is

$$\mathcal{D}_{st}\sigma(k_1,\cdots,k_p)=w_{st}(Ax_tx_s'A+(x_t'Ax_s)A)+\delta^{st}\sum_{\ell=1}^n w_{sv}Ax_vx_v'A. \tag{6}$$

Thus $w_{st}=0$ implies

$$\mathcal{D}_{st}\sigma(x_1,\cdots,x_p)=\delta^{st}\sum_{v=1}^p w_{sv}Ax_vx_v'A=\delta^{st}AXW_sX'A, \tag{7}$$

with W_s a diagonal matrix with column s of W in the diagonal. For an $s\neq t$ with $w_{st}=0$ we have $\mathcal{D}_{st}\sigma(x_1,\cdots,x_p)=0$.

There is R code in the appendix implementing formulas (5) and (6), as well as code for checking the formulas numerically using numDeriv (Gilbert and Varadhan (2019)).

4 Jacobi

Following Jacobi (1846) we build up K using elementary rotations, constructed by using the unit vectors e_i, which have element i equal to one and all other elements equal to zero. Suppose $T_{ij}(x)$ is a matrix with column t_i equal to $e_i\sin x+e_j\cos x$ and column x_j equal to $e_j\sin x-e_i\cos x$, where e_i and e_j are units vectors. Column k for $k\notin\{i,j\}$ is equal to e_k. More explicitly we could write X as $X_{ij}(\alpha,\beta)$. Clearly X is square orthonormal.

The general idea of the Jacobi method is that we have an infinite sequence $(i(\nu),j(\nu))$ of *pivots*, leading to the infinite sequence $X^\nu_{i(\nu),j(\nu)}(\alpha^\nu,\beta^\nu)$ where &□^{□}$ and β^ν are chosen to minimize

σ. We then replace $A^{(\nu)}$ by $A^{(\nu+1)} =$ and $X^{(\nu)}$ by $X^{(\nu+1)}$ $\overline{X}$ and go to the next pivot in the sequence.

$$X^{(\nu+1)} = X^{(\nu)} T^{\nu}_{i(\nu),j(\nu)}(\alpha^{(\nu)}, \beta^{(\nu)})$$

$$(\alpha^{(\nu)}, \beta^{(\nu)}) = \underset{\alpha^2+\beta^2=1}{\text{argmin}} \ \sigma()$$

Let's look at the problem of optimizing , i.e. making a single pivot. Then σ is a function of (α, β) on the unit cicle. Define the symmetric matrix $\overline{A} = X'AX$. Then $\overline{a}_{kl} = x_k'Ax_l$, which means that for $k \neq i$ and $k \neq j$ as well as $l \neq i$ and $l \neq j$ we have $\overline{a}_{kl} = a_{kl}$. For $k \notin \{i, j\}$ we have

$$\overline{a}_{ik} = x_i'Ae_k = x_i'a_k = \alpha a_{ik} + \beta a_{jk}, \tag{8}$$
$$\overline{a}_{jk} = x_j'Ae_k = x_j diagonal' a_k = -\beta a_{ik} + \alpha a_{jk}. \tag{9}$$

Moreover

$$\overline{a}_{ij} = x_i'Ax_j = (\alpha e_i + \beta e_j)'A(\alpha e_j - \beta e_i) = (\alpha^2 - \beta^2)a_{ij} + \alpha\beta(a_{jj} - a_{ii}), \tag{10}$$
$$\overline{a}_{ii} = x_i'Ax_i = (\alpha e_i + \beta e_j)'A(\alpha e_i + \beta e_j) = \alpha^2 a_{ii} + \beta^2 a_{jj} + 2\alpha\beta a_{ij}, \tag{11}$$
$$\overline{a}_{jj} = x_j'Ax_j = (-\beta e_i + \alpha e_j)'A(-\beta e_i + \alpha e_j) = \beta^2 a_{ii} + \alpha^2 a_{jj} - 2\alpha\beta a_{ij}. \tag{12}$$

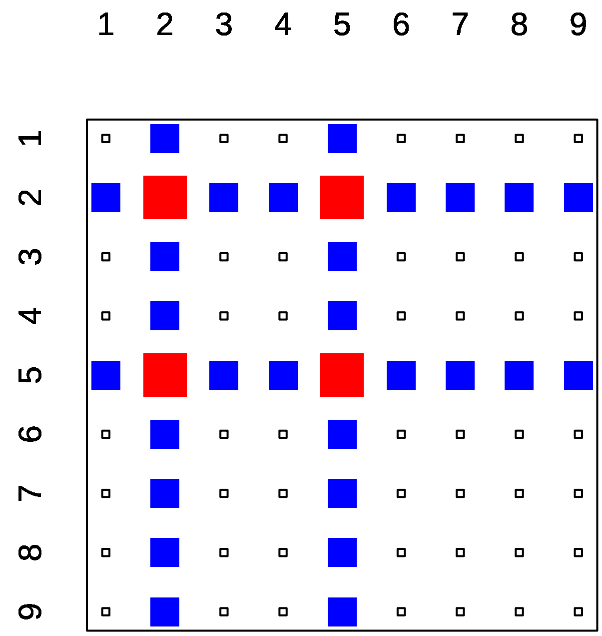
In summary

$$\overline{a}_{kl} = \begin{cases} a_{kl} & \text{if } k \neq \{i,j\} \text{ and } l \neq \{i,j\}, \\ & \text{if } k = i \text{ and } l \neq \{i,j\} \text{ or } k \neq \{i,j\} \text{ and } l = i, \\ & \text{if } k = j \text{ and } l \neq \{i,j\}, \\ & \text{if } k = i \text{ and } l = j, \\ & \text{if } k = i \text{ and } l = i, \\ & \text{if } k = j \text{ and } l = j. \end{cases} \tag{13}$$

Thus

$$\sigma(X) = \sum_{k \notin \{i,j\}}^{n} \sum_{l \notin \{i,j\}}^{n} w_{kl} a_{kl}^2 + \tag{14}$$
$$+ 2 \sum_{k \notin \{i,j\}}^{n} w_{ik}(\alpha a_{ik} + \beta a_{jk})^2 + \tag{15}$$
$$+ 2 \sum_{k \notin \{i,j\}}^{n} w_{jk}(-\beta a_{ik} + \alpha a_{jk})^2 + \tag{16}$$
$$+ 2w_{ij}\{(\alpha^2 - \beta^2)a_{ij} + \alpha\beta(a_{jj} - a_{ii})\}^2 + \tag{17}$$
$$+ w_{ii}(\alpha^2 a_{ii} + \beta^2 a_{jj} + 2\alpha\beta a_{ij})^2 + \tag{18}$$
$$+ w_{jj}(\beta^2 a_{ii} + \alpha^2 a_{jj} - 2\alpha\beta a_{ij})^2 \tag{19}$$

```
par(pty="s")
jacobiPlot(2, 5, 9)
```



Trigonometry

$\alpha = \sin(\theta)$ and $\beta = \cos(\theta)$

5 The Sequence

The *strategy*

6 Majorization

$$\sigma(y_1,\cdots,y_p)=\sigma(x_1+(y_1-x_1),\cdots,x_p+(y_p-x_p))\leq \sigma(x_1,\cdots,x_p)+\sum_{s=1}^p(y_s-x_s)'\mathcal{D}_s\sigma(x_1,\cdots,x_p)+ \frac{1}{2}\max_{0\leq\theta\leq 1}\sum_{s=1}^p\sum_{t=1}^p(y_s-x_s)'\{\mathcal{D}_{st}\sigma(x_1+\theta(x_1-y_1),\cdots,x_p+\theta(x_p-y_p))\}(y_t-x_t). \quad (20)$$

7 Applications

7.1 Symmetric Matrices

All eigenvalues Some eigenvalues

7.2 Pairs of Matrices

7.3 Rectangular Matrices

7.4 Simultaneous Diagonalization

7.5 DMCA

7.6 GCCA

8 Appendix: Code

8.1 pattern.R

```
mPrint <- function(x,
                   digits = 6,
                   width = 8,
                   format = "f",
                   flag = "+") {
  print(noquote(
    formatC(
      x,
      digits = digits,
      width = width,
      format = format,
      flag = flag
    )
  ))
}

butLast <- function(x, m = 1) {
  return(rev(rev(x)[-(1:m)]))
}

butFirst <- function(x, m = 1) {
  return(x[-(1:m)])
}
```

References

De Leeuw, J. 2017. "Jacobi Eigen in R/C with Lower Triangular Column-wise Compact Storage." 2017. https://jansweb.netlify.app/publication/deleeuw-e-17-o/deleeuw-e-17-o.pdf.

———. 2018. "Simultaneous Diagonalization in R/C." 2018. https://jansweb.netlify.app/publication/deleeuw-e-18-g/deleeuw-e-18-g.pdf.

De Leeuw, J., and D. B. Ferrari. 2008. “Using Jacobi Plane Rotations in R.” Preprint Series 556. Los Angeles, CA: UCLA Department of Statistics. https://jansweb.netlify.app/publication/deleeuw-ferrari-r-08-a/deleeuw-ferrari-r-08-a.pdf.

De Leeuw, J., and S. Pruzansky. 1978. “A New Computational Method to Fit the Weighted Euclidean Distance Model.” *Psychometrika* 43: 479–90.

Gilbert, P., and R. Varadhan. 2019. *numDeriv: Accurate Numerical Derivatives*. https://CRAN.R-project.org/package=numDeriv.

Jacobi, C. G. J. 1846. “Über ein leichtes Verfahren die in der Theorie der Secülarstörungen vorkommenden Gleichungen numerisch aufzulösen.” *Journal für Die Reine Und Angewandte Mathematik (Crelle’s Journal)* 30: 51–94.