# BAYES LINEAR ANALYSIS

[This article appears in the Encyclopaedia of Statistical Sciences, Update volume 3, 1998, Wiley.]

The Bayes linear approach is concerned with problems in which we want to combine prior judgements of uncertainty with observational data, and we use EXPECTATION rather than probability as the primitive for expressing these judgements. This distinction is of particular relevance in complex problems with too many sources of information for us to be comfortable in making a meaningful full joint prior probability specification of the type required for a BAYESIAN ANALYSIS. Therefore, we seek methods of prior specification and analysis which do not require this extreme level of detail. For such problems, expectation may be considered as a more natural primitive than probability; see PREVISION for a summary of de Finetti's treatment of expectation as a primitive, and from a rather different viewpoint, see Whittle [15]. Thus, the Bayes linear approach is similar in spirit to a full Bayes analysis, but is based on a simpler approach to prior specification and analysis, and so offers a practical methodology for analysing partially specified beliefs for large problems.

## ADJUSTED MEANS AND VARIANCES

In the Bayes linear approach, we make direct prior specifications for that collection of means, variances and covariances which we are both willing and able to assess, and update these prior assessments by linear fitting. Suppose that we have two collections of random quantities, namely vectors $\boldsymbol{B} = (B_1, ..., B_r)$, $\boldsymbol{D} = (D_0, D_1, ..., D_s)$, where $D_0 = 1$, and we intend to observe $\boldsymbol{D}$ in order to improve our assessments of belief over $\boldsymbol{B}$. The *adjusted* or *Bayes linear* expectation for $B_i$ given $\boldsymbol{D}$ is the linear combination $\boldsymbol{a}_i^T\boldsymbol{D}$ minimising $\mathrm{E}((B_i - \boldsymbol{a}_i^T\boldsymbol{D})^2)$ over choices of $\boldsymbol{a}_i$. To do so, we must specify prior mean vectors and variance matrices for $\boldsymbol{B}$ and $\boldsymbol{D}$ and a covariance matrix between $\boldsymbol{B}$ and $\boldsymbol{D}$. The adjusted expectation vector, $\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})$, for $\boldsymbol{B}$ given $\boldsymbol{D}$, is evaluated as

$$\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B}) = \mathrm{E}(\boldsymbol{B}) + \mathrm{cov}(\boldsymbol{B}, \boldsymbol{D})(\mathrm{var}(\boldsymbol{D}))^{-1}(\boldsymbol{D} - \mathrm{E}(\boldsymbol{D})) \tag{1}$$

If $\mathrm{var}(\boldsymbol{D})$ is not invertible, then we may use an appropriate generalised inverse in the above, and following, equations. The *adjusted variance matrix* for $\boldsymbol{B}$ given $\boldsymbol{D}$, denoted by $\mathrm{var}_{\boldsymbol{D}}(\boldsymbol{B})$, is evaluated as

$$\mathrm{var}_{\boldsymbol{D}}(\boldsymbol{B}) = \mathrm{var}(\boldsymbol{B} - \mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})) = \mathrm{var}(\boldsymbol{B}) - \mathrm{cov}(\boldsymbol{B}, \boldsymbol{D})(\mathrm{var}(\boldsymbol{D}))^{-1}\mathrm{cov}(\boldsymbol{D}, \boldsymbol{B}) \qquad (2)$$

Stone [14], and Hartigan [9] are among the first to discuss the role of such assessments in Bayes analysis with partial prior specification. For examples of papers concerned with the practical details of assessing and generalising Bayes linear estimators in particular problems, see Mouchart and Simar [10], O'Hagan [11], and Cocchi and Mouchart [1].

We may write $\boldsymbol{B}$ as the sum of the two uncorrelated quantities $(\boldsymbol{B} - \mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B}))$ and $\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})$, so that $\mathrm{var}(\boldsymbol{B}) = \mathrm{var}(\boldsymbol{B} - \mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})) + \mathrm{var}(\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B}))$. We term $\mathrm{rvar}_{\boldsymbol{D}}(\boldsymbol{B}) = \mathrm{var}(\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B}))$ the *resolved variance matrix*, so that $\mathrm{var}(\boldsymbol{B}) = \mathrm{var}_{\boldsymbol{D}}(\boldsymbol{B}) + \mathrm{rvar}_{\boldsymbol{D}}(\boldsymbol{B})$, and informally the resolved variance matrix expresses the uncertainty about $\boldsymbol{B}$ removed by the adjustment.

## INTERPRETATIONS OF BELIEF ADJUSTMENT

(1) Within the usual Bayesian view, adjusted expectation offers a simple, tractable approximation to conditional expectation, which is useful in complex problems, while adjusted variance is a strict upper bound to expected posterior variance, over all prior specifications consistent with the given moment structure. The approximations are exact in certain important special cases, and in particular if the joint probability distribution of $\boldsymbol{B}, \boldsymbol{D}$ is multivariate normal. Therefore, there are strong formal relationships between Bayes linear calculations and the analysis of Gaussian structures, so that the linear adjustments arise in contexts as diverse as dynamic linear models and KRIGING.

(2) $\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})$ may be viewed as an estimator of $\boldsymbol{B}$, combining the data with simple aspects of prior beliefs in an intuitively plausible manner, so that $\mathrm{var}_{\boldsymbol{D}}(\boldsymbol{B})$ is the expected mean square error of $\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})$. As a class of estimators, the Bayes linear rules have certain important admissibility properties; see LINEAR ESTIMATORS, BAYES.

(3) Adjusted expectation is numerically equivalent to conditional expectation in the particular case where $\boldsymbol{D}$ comprises the indicator functions for the elements of a partition, i.e. where each $D_i$ takes value one or zero and precisely one element $D_i$ will equal one. We may view adjusted expectation as a natural generalisation of the approach to conditional expectation based on 'conditional' quadratic penalties (*see* PREVISION) where we drop the restriction that we may only condition on the indicator functions for a partition. Here, adjusted variance may be interpreted as a primitive quantity, analogous to prior variance but

applied to the residual variation when we have extracted the variation in $\boldsymbol{B}$ accounted for by adjustment on $\boldsymbol{D}$.

(4) A more fundamental interpretation, which subsumes each of the above views, is based on foundational consideration concerning the implications of a partial collection of prior belief statements about $\boldsymbol{B}, \boldsymbol{D}$ for the posterior assessment that we may make for the expectation of $\boldsymbol{B}$ having observed $\boldsymbol{D}$. Any linkage between belief statements at different times requires some form of temporal COHERENCE condition. The temporal sure preference condition says, informally, that if it is logically necessary that you will prefer a certain small random penalty $A$ to $C$ at some given future time, then you should not now have a strict preference for penalty $C$ over $A$. This condition is weak enough to be acceptable as a temporal coherence condition for many situations, and has the consequence that your actual posterior expectation, $\mathrm{E}_T(\boldsymbol{B})$, at time $T$ when you have observed $\boldsymbol{D}$, satisfies the relation

$$\mathrm{E}_T(\boldsymbol{B}) = \mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B}) + \boldsymbol{R}, \tag{3}$$

where $\boldsymbol{R}$ has, a priori, zero expectation and is uncorrelated with $\boldsymbol{D}$. Therefore, adjusted expectation may be viewed as a prior inference for your actual posterior judgements, which resolves a portion of your current variance for $\boldsymbol{B}$, and whose difference from the posterior judgement is not linearly predictable. The larger this resolved variance, the more useful is the prior analysis, and this is determined by the choice of $\boldsymbol{D}$. In the special case where $\boldsymbol{D}$ represents a partition, $\mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B})$ is equal to the conditional expectation given $\boldsymbol{D}$, and $\boldsymbol{R}$ has conditional expectation zero for each member of the partition. In this view, relationships between actual belief revisions and formal analysis based on partial prior specifications are entirely derived through stochastic relations such as (3); see the discussion in Goldstein [7].

(5) The geometric interpretation of adjusted beliefs is as follows. For any collection $\boldsymbol{C} = (C_1, C_2, ...)$ of random quantities, we denote by $\langle \boldsymbol{C} \rangle$ the collection of (finite) linear combinations $\sum_i r_i C_i$ of the elements of $\boldsymbol{C}$. Adjusted expectation is linear, that is $\mathrm{E}_{\boldsymbol{D}}(X+Y) = \mathrm{E}_{\boldsymbol{D}}(X) + \mathrm{E}_{\boldsymbol{D}}(Y)$, so that defining adjusted expectation over the elements of $\boldsymbol{C}$ is equivalent to defining adjusted expectation over $\langle \boldsymbol{C} \rangle$. We view $\langle \boldsymbol{C} \rangle$ as a vector space. Prior covariance acts as an inner product on this space. If we choose $\boldsymbol{C}$ to be the union of the collection of elements of the vectors $\boldsymbol{B}$ and $\boldsymbol{D}$, then the adjusted expectation of an element

$Y \in \langle \boldsymbol{B} \rangle$ given $\boldsymbol{D}$ is the orthogonal projection of $Y$ into the linear subspace $\langle \boldsymbol{D} \rangle$, and adjusted variance is the squared distance between $Y$ and $\langle \boldsymbol{D} \rangle$. Each of the finite-dimensional analyses described in this article has an infinite-dimensional counterpart within this construction. For example, the usual Bayes formalism is represented as follows. Corresponding to the vector space $\langle C \rangle$ is the collection of all random variables defined over the outcome space, and the inner product space may therefore be identified with the Hilbert space of square integrable functions over the outcome space, with respect to the joint prior measure over the outcomes, with squared norm the expected squared distance between functions, so that conditional expectation given a sample is equivalent to orthogonal projection into the subspace of all functions defined over the sample space. This Hilbert space, for many problems, is large and hard to specify. Bayes linear analysis may be viewed as restricting prior specification and subsequent projection into the largest subspace of the full space that we are prepared to specify prior beliefs over. The geometric formulation extends the Bayes linear approach to general analyses of uncertainty over linear spaces; see, for example, Wilkinson and Goldstein [16], in which a Bayes linear approach is developed for adjusting beliefs over variance matrices considered as elements of an appropriate inner product space.

## INTERPRETIVE AND DIAGNOSTIC MEASURES

Much of Bayes linear methodology is built around the following interpretive and diagnostic cycle: (i) we interpret the expected effects of the adjustment, a priori; (ii) given observations, we interpret the outcome of the actual adjustment; (iii) we make diagnostic comparisons between observed and expected beliefs. These comparisons may be carried out as follows.

(i) The expected effects of the adjustment of vector $\boldsymbol{B}$ by $\boldsymbol{D}$ may be examined through the eigenstructure of the *resolution transform*, which is defined to be $\boldsymbol{T}_{\boldsymbol{D}} = (\mathrm{var}(\boldsymbol{B}))^{-1}\mathrm{rvar}_{\boldsymbol{D}}(\boldsymbol{B})$. Denote the eigenvectors of $\boldsymbol{T}_{\boldsymbol{D}}$ as $\boldsymbol{c}_1, ...\boldsymbol{c}_r$, corresponding to eigenvalues $1 \geq \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_r \geq 0$, and let $Z_i = \boldsymbol{c}_i^T \boldsymbol{B}$, where each $\boldsymbol{c}_i$ is scaled so that $\mathrm{var}(Z_i) = 1$. By analogy with CANONICAL ANALYSIS, we call the collection $Z_1, ..., Z_r$ the *canonical variables* for the belief adjustment. The eigenvalues are termed the *canonical resolutions* as, for each $i$, $\lambda_i = \mathrm{rvar}_{\boldsymbol{D}}(Z_i)$. The canonical variables form a grid of uncorrelated directions over $\langle \boldsymbol{B} \rangle$ which summarise the effects of the adjustment in the following sense. For any $Y = \boldsymbol{a}^T \boldsymbol{B}$, we have $\mathrm{rvar}_{\boldsymbol{D}}(Y) = \mathrm{cov}(Y, \boldsymbol{T}_{\boldsymbol{D}} Y)$, from which we may deduce that, for any $Y \in \langle \boldsymbol{B} \rangle$, with prior variance 1, $\text{rvar}_{\boldsymbol{D}}(Y) = \sum_i \lambda_i \text{cov}(Y, Z_i)$, and, in particular, for each $i$, $\text{rvar}_{\boldsymbol{D}}(Z_i)$ maximises $\text{rvar}_{\boldsymbol{D}}(\boldsymbol{a}^T\boldsymbol{B})$ over all choices $\boldsymbol{a}$ for which $\boldsymbol{a}^T\boldsymbol{B}$ is uncorrelated with $Z_1, ..., Z_{i-1}$. Therefore, adjustment by $\boldsymbol{D}$ is mainly informative about those directions in $\langle\boldsymbol{B}\rangle$ with large correlations with canonical variables having large canonical resolutions; see the discussion in Goldstein [3]. Comparisons of canonical structures therefore provide guidance, for example in choosing between alternative choices of sampling frame or experimental design.

(ii) We may summarise the actual adjustment in beliefs as follows. Observing the vector $\boldsymbol{D} = \boldsymbol{d}$ gives, for any $Y \in \langle\boldsymbol{B}\rangle$, observed value $\text{E}_{\boldsymbol{d}}(Y)$ for the adjusted expectation $\text{E}_{\boldsymbol{D}}(Y)$. We construct the *bearing* of the adjustment from the canonical variables as $Z_{\boldsymbol{d}} = \sum_i \text{E}_{\boldsymbol{d}}(Z_i)Z_i$. We have $\text{cov}(Y, Z_{\boldsymbol{d}}) = \text{E}_{\boldsymbol{d}}(Y) - \text{E}(Y)$. Therefore (a) for any $Y$ uncorrelated with $Z_{\boldsymbol{d}}$, the prior and adjusted expectations are the same; (b) $\text{var}(Z_{\boldsymbol{d}})$ maximises over $Y \in \langle\boldsymbol{B}\rangle$ the value of $S_{\boldsymbol{d}}(Y) = (\text{E}_{\boldsymbol{d}}(Y) - \text{E}(Y))^2/\text{var}(Y)$, this maximum occurring for $Y = Z_{\boldsymbol{d}}$; (c) multiplying $Z_{\boldsymbol{d}}$ by a constant multiplies each adjustment $(\text{E}_{\boldsymbol{d}}(Y) - \text{E}(Y))$ by that constant. Thus, all changes in belief are in the 'direction' of $Z_{\boldsymbol{d}}$, and $\text{var}(Z_{\boldsymbol{d}})$ represents the 'magnitude' of the adjustment in belief; for discussion and applications, see Goldstein [5]. In the infinite-dimensional version of the analysis, the bearing is constructed as the Riesz representation of the adjusted expectation functional on the inner product space over $\langle\boldsymbol{B}\rangle$, and in a full Bayes analysis, the bearing is usually equivalent to a normalised version of the likelihood.

(iii) A natural diagnostic comparison between observed and expected adjustments is to compare the maximum value of $S_{\boldsymbol{d}}(Y)$, namely $\text{var}(Z_{\boldsymbol{d}})$, with the prior expectation for this maximum. This prior expectation is equal to the trace of the resolution transform, $\boldsymbol{T}_{\boldsymbol{D}}$, for the adjustment. Thus, if $\text{var}(Z_{\boldsymbol{d}})$ is much larger than the trace then this may suggest that we have formed new beliefs which are surprisingly discordant with our prior judgements. It is important to identify such discrepancies, but whether this causes us to re-examine the prior specification, or the adjusted assessments, or the data, or to accept the analysis depends entirely on the context.

Each of the above interpretive and diagnostics quantities has a corresponding partial form to assess the effect on adjusted beliefs of individual aspects of a collection of pieces of evidence. In the simplest case, suppose that we split data $\boldsymbol{D}$ into portions $\boldsymbol{E}$ and $\boldsymbol{F}$, based on some criterion such as time or place of measurement. The eigenstructure of $T_{\boldsymbol{E}}$ summarises the usefulness of $\boldsymbol{E}$ for the assessment of $\boldsymbol{B}$. Similarly, the eigenstructure of the partial resolution matrix $T_{[\boldsymbol{F}/\boldsymbol{E}]} = T_{\boldsymbol{D}} - T_{\boldsymbol{E}}$ summarises the additional effect on $\boldsymbol{B}$ of adding the adjustment by $\boldsymbol{F}$ to that of $\boldsymbol{E}$, and the trace of $T_{[\boldsymbol{F}/\boldsymbol{E}]}$ is equal to the prior expectation for the maximum value of $(\mathrm{E}_{\boldsymbol{d}}(Y) - \mathrm{E}_{\boldsymbol{e}}(Y))^2/\mathrm{var}(Y)$, for $Y \in \langle \boldsymbol{B} \rangle$. The partial bearing $Z_{[\boldsymbol{f}/\boldsymbol{e}]} = Z_{\boldsymbol{d}} - Z_{\boldsymbol{e}}$ summarises the additional changes in adjusted expectation from observing $\boldsymbol{F} = \boldsymbol{f}$. We term $\mathrm{corr}(Z_{\boldsymbol{e}}, Z_{[\boldsymbol{f}/\boldsymbol{e}]})$ the *path correlation* and this quantity is a measure of the degree of support, if positive, or conflict, if negative, between the two collections of evidence in determining the overall adjustment of beliefs.

Each feature of the above analysis may be usefully displayed using *Bayes linear influence diagrams*. These diagrams are constructed in a similar manner to standard influence diagrams, but under the relation that the vectors $\boldsymbol{B}$ and $\boldsymbol{C}$ are *separated* by the vector $\boldsymbol{D}$ provided that $\mathrm{corr}(\boldsymbol{B} - \mathrm{E}_{\boldsymbol{D}}(\boldsymbol{B}), \boldsymbol{C} - \mathrm{E}_{\boldsymbol{D}}(\boldsymbol{C})) = 0$. Separation acts as a generalised conditional independence property (see Smith [13]), and diagrams defined over vector nodes based upon such a definition may be manipulated by the usual rules governing influence diagrams. Because these diagrams are defined through the covariance structure, they share many formal properties with Gaussian diagrams. These diagrams may be used firstly to build up the qualitative form of the covariance structure between the various components of the problem, and secondly to give a simple graphical representation of the interpretive and diagnostic features that we have described. For example, we may shade the outer ring of a node to express the amount of variation that is reduced by partial adjustment by each parent node, and we may shade within a node to show diagnostic warnings of differences between observed and expected outcomes; for details and computational implementation see Goldstein and Wooff [8]

## SECOND ORDER EXCHANGEABILITY

One of the principle motivations for the Bayes linear approach is to develop a methodology which is based strictly on the combination of meaningful prior judgements with observational data. Central to this aim is the need for an approach to statistical modelling in which models may be constructed directly from simple collections of judgements over observable quantities. We achieve this direct construction of statistical models using the representation theorem for second-order exchangeable random vectors.

We say that an infinite sequence of vectors $\boldsymbol{X}_i$ is *second-order exchangeable* if the mean, variance and covariance structure is invariant under permutation, namely $\mathrm{E}(\boldsymbol{X}_i) = \mu$, $\mathrm{var}(\boldsymbol{X}_i) = \Sigma$, $\mathrm{cov}(\boldsymbol{X}_i, \boldsymbol{X}_j) = \Gamma, \forall i \neq j$. We may represent each $\boldsymbol{X}_i$ as the uncorrelated sum of an underlying 'population mean' $\boldsymbol{M}$, and individual variation $\boldsymbol{R}_i$, according to the following representation theorem: for each $i$, we may write $\boldsymbol{X}_i$ as $\boldsymbol{X}_i = \boldsymbol{M} + \boldsymbol{R}_i$, where $\mathrm{E}(\boldsymbol{M}) = \mu, \mathrm{var}(\boldsymbol{M}) = \Gamma, \mathrm{E}(\boldsymbol{R}_i) = 0, \mathrm{var}(\boldsymbol{R}_i) = \Sigma - \Gamma, \forall i$ and the vectors $\boldsymbol{M}, \boldsymbol{R}_1, \boldsymbol{R}_2, \ldots$ are mutually uncorrelated; see Goldstein [4]. This representation is similar in spirit to de Finetti's representation theorem for fully exchangeable sequences; see EXCHANGEABILITY. However, while second order exchangeability is usually straightforward to specify over the observable quantities, the specification of beliefs for a fully exchangeable sequence requires such an extreme level of detail of specification that, even for the simplest of problems, it is impossible to apply the full representation theorem in practice to create statistical models from beliefs specified over observables.

Suppose that we want to adjust predictive beliefs about a future observation $\boldsymbol{X}_j$ given a sample of $n$ previous observations $\boldsymbol{X}_{[n]} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$. We obtain the same results if we directly adjust $\boldsymbol{X}_j$ by $\boldsymbol{X}_{[n]}$ or if we adjust $\boldsymbol{M}$ by $\boldsymbol{X}_{[n]}$ and derive adjusted beliefs for $\boldsymbol{X}_j$ via the representation theorem; see Goldstein [6]. This adjustment is of a simple form in the following sense. The canonical variables for the adjustment of $\boldsymbol{M}$ by $\boldsymbol{X}_{[n]}$ are the same for each sample size $n$, and if the canonical resolution for a particular canonical variable is $\lambda$, for $n = 1$, then the corresponding canonical resolution for a sample of $n$ is $n\lambda/(1 + (n-1)\lambda)$. Therefore, the qualitative features of the adjustment are the same for all sample sizes and it is simple and natural to compare choices of possible sample sizes based on analysis of the eigenstructure of the resolution transform.

## CONCLUDING COMMENTS

We have discussed general features which characterise the Bayes linear approach. For examples of the application of the methodology, see Craig et al [2] and O'Hagan et al [12], which illustrate how the Bayes linear approach may be used to combine expert assessments with observational information for large and complex problems in which it would be extremely difficult to develop full Bayes solutions.

# References

[1] Cocchi, D. and Mouchart, M. (1996). Quasi-linear Bayes estimation in stratified finite populations, *J. R. Statist. Soc. B*, **58**, 293-300.

[2] Craig, P. S., Goldstein, M., Seheult, A. H. and Smith, J. A. (1996). Bayes linear strategies for history matching of hydrocarbon reservoirs, in *Bayesian Statistics 5*, Bernardo, J.M. et al eds, 69-98, Oxford University Press.

[3] Goldstein, M (1981 ). Revising previsions: a geometric interpretation (with discussion)*J. R. Statist. Soc. B*, **43** 105–130

[4] Goldstein, M. (1986). Exchangeable belief structures *J. Amer. Statist. Ass.*, **81**, 971-976.

[5] Goldstein, M. (1988). The Data Trajectory, *Bayesian Statistics 3*, Bernardo, J.M. et al. eds, 189-209, Oxford University Press.

[6] Goldstein, M. (1994 ). Revising exchangeable beliefs: subjectivist foundations for the inductive argument, in *Aspects of Uncertainty: A Tribute to D. V. Lindley*, Freeman, P. and Smith, A.F.M. eds, 201-222, Wiley.

[7] Goldstein, M. (1996). Prior inferences for posterior judgements, in *Proceedings of the 10th International Congress of Logic, Methodology and Philosophy of Science* , M. L. D. Chiara et al eds, Kluwer.

[8] Goldstein, M. and Wooff, D. A. (1995). Bayes linear computation: concepts, implementation and programming environment. *Statistics and Computing*, **5** 327-341.

[9] Hartigan, J. A. (1969) Linear Bayes methods. *J. Roy. Statist. Soc., B*, **31** 446-454.

[10] Mouchart, M. and Simar, L. (1980) Least squares approximation in Bayesian analysis, in *Bayesian Statistics*, Bernardo, J.M. et al eds, University Press, Valencia, 207-222.

[11] O'Hagan, A. (1987) Bayes linear estimators for randomized response models, *J. Amer. Statist. Assoc.*, **82** 580-585.

[12] O'Hagan, A. and Glennie, E. B. and Beardsall, R. E. 1992. Subjective modelling and Bayes linear estimation in the UK water industry. *Applied Statistics*, **41**, 563–577.

[13] Smith, J. Q. (1990 ). Statistical principles on graphs, in *Influence Diagrams, Belief Nets and Decision Analysis*, Smith, J. Q. and Oliver, R. M., eds.,Wiley, Chichester.

[14] Stone, M. (1963) Robustness of non-ideal decision procedures. *J. Amer. Statist. Ass.*, **58** 480-486.

[15] Whittle, P. (1992) *Probability via expectation*, Springer Verlag.

[16] Wilkinson, D.J. and Goldstein, M. (1996) Bayes linear adjustment for variance matrices, in *Bayesian Statistics 5*, Bernardo, J.M. et al eds, 791-800, Oxford University Press.

(BAYESIAN ANALYSIS
CANONICAL ANALYSIS
COHERENCE
EXCHANGEABILITY
EXPECTATION
LINEAR ESTIMATORS, BAYES
PREVISION)

MICHAEL GOLDSTEIN