The complex multivariate Gaussian distribuion

Robin K. S. Hankin
Auckland University of Technology

Abstract

Here, **cmvnorm**, a complex generalization of the **mvtnorm** package is presented. An application in the context of a complex Gaussian process as fitted to the Weierstrass sigma function is given

Keywords: Complex multivariate Gaussian distribution, Gaussian process, Weierstrass sigma function, emulator .

1. Introduction

The multivariate Gaussian distribution is well supported in R (R Core Team 2014; Genz, Bretz, Miwa, Mi, Leisch, Scheipl, and Hothorn 2014), having density function

$$f\left(\mathbf{x};\boldsymbol{\mu},\Sigma\right)=\frac{e^{-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}\right)^T\Sigma^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)}}{\sqrt{\left|2\pi\Sigma^{-1}\right|}}\qquad\mathbf{x}\in\mathbb{R}^n. \tag{1}$$

Here, $\boldsymbol{\mu}=\mathbb{E}\mathbf{x}\in\mathbb{R}^n$ is the mean vector and $\Sigma=\mathbb{E}\left(\mathbf{x}-\boldsymbol{\mu}\right)\left(\mathbf{x}-\boldsymbol{\mu}\right)^T$ the variance of random variable $\mathbf{X}$; we write $\mathbf{X}\sim\mathcal{N}\left(\boldsymbol{\mu},\Sigma\right)$. One natural generalization would be to consider $\mathbf{Z}\sim\mathcal{NC}\left(\boldsymbol{\mu},\Gamma\right)$, the complex multivariate Gaussian, with density function

$$f\left(\mathbf{z};\boldsymbol{\mu},\Gamma\right)=\frac{e^{-\left(\mathbf{z}-\boldsymbol{\mu}\right)^\star\Gamma^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)}}{\left|\pi\Gamma^{-1}\right|}\qquad\mathbf{z}\in\mathbb{C}^n \tag{2}$$

where $\mathbf{z}^\star$ denotes the Hermitian transpose of complex vector $\mathbf{z}$. Now $\boldsymbol{\mu}\in\mathbb{C}^n$ is the complex mean and $\Gamma=\mathbb{E}\left(\mathbf{Z}-\boldsymbol{\mu}\right)\left(\mathbf{Z}-\boldsymbol{\mu}\right)^\star$ is the complex variance; Γ is a Hermitian positive definite matrix. Note the simpler form of equation 2, essentially due to Gauss's integral operating more cleanly over the complex plane than the real line:

$$\int_{\mathbb{C}}e^{-z^\star z}\,dz=\int_{x\in\mathbb{R}}\int_{y\in\mathbb{R}}e^{-\left(x^2+y^2\right)}\,dx\,dy=\int_{\theta=0}^{2\pi}\int_{r=0}^{\infty}e^{-r^2}r\,dr\,d\theta=\pi.$$

A zero mean complex random vector $\mathbf{Z}$ is said to be *circularly symmetric* (Goodman 1963) if $\mathbb{E}\mathbf{Z}\mathbf{Z}^T=\mathbf{0}$, or equivalently $\mathbf{Z}$ and $e^{i\alpha}\mathbf{Z}$ have identical distributions for any $\alpha\in\mathbb{R}$. Equation 2 clearly has this property.

Most results from real multivariate analysis have a direct generalization to the complex case, as long as "transpose" is replaced by "Hermitian transpose". For example, $\mathbf{X}\sim\mathcal{N}\left(\mathbf{0},\Sigma\right)$ implies $B\mathbf{X}\sim\mathcal{N}\left(\mathbf{0},B^T\Sigma B\right)$ for any constant matrix B, and analogously $\mathbf{Z}\sim\mathcal{NC}\left(\mathbf{0},\Gamma\right)$

implies $B\mathbf{Z} \sim \mathcal{NC}\left(\mathbf{0}, B^{\star}\Gamma B\right)$. Similar generalizations operate for Schur complement methods on partitioned matrices.

Also, linear regression works without modification. Specifically, consider $\mathbf{y} \in \mathbb{R}^n$. If $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where X is a $n \times p$ design matrix, $\boldsymbol{\beta} \in \mathbb{R}^p$ a vector of regression coefficients and $\boldsymbol{\epsilon} \sim \mathcal{N}\left(0, \sigma^2 A\right)$ is a vector of errors, then $\hat{\boldsymbol{\beta}} = \left(X^T A^{-1} X\right) X^T A^{-1}\mathbf{y}$ is the maximum likelihood estimator for $\boldsymbol{\beta}$. The complex generalization is $\hat{\boldsymbol{\beta}} = \left(X^{\star} A^{-1} X\right)^{-1} X^{\star} A^{-1}\mathbf{z}$, where A itself may be complex.

This short vignette introduces the **cmvnorm** package which furnishes some functionality for the complex multivariate Gaussian distribution, and applies it in the context of a complex generalization of the **emulator** package.

2. The package in use

Random complex vectors are generated using the `rcmvnorm()` function, analogous to `rmvnorm()`:

```
> set.seed(1)
> require(cmvnorm,quietly=TRUE)
> cm <- c(1,1i)
> cv <- matrix(c(2,1i,-1i,2),2,2)
> (z <- rcmvnorm(6, mean=cm, sigma=cv))

                      [,1]                    [,2]
[1,] 0.9680986+0.5525419i  0.0165969+2.9770976i
[2,] 0.2044744-1.4994889i  1.8320765+0.8271259i
[3,] 1.0739973+0.2279914i -0.7967020+0.1736071i
[4,] 1.3171073-0.9843313i  0.9257146+0.5524913i
[5,] 1.3537303-0.8086236i -0.0571337+0.3935375i
[6,] 2.9751506-0.1729231i  0.3958585+3.3128439i
```

Function `dcmvnorm()` returns the density according to equation 2:

```
> dcmvnorm(z,cm,cv)

[1] 5.103754e-04 1.809636e-05 2.981718e-03 1.172242e-03 4.466836e-03
[6] 6.803356e-07
```

So it is possible to determine a maximum likelihood for the mean using direct numerical optimization

```
> helper <- function(x){c(x[1]+1i*x[2], x[3]+1i*x[4])}
> objective <- function(x){-sum(dcmvnorm(z,mean=helper(x),sigma=cv,log=TRUE))}
> optim(c(1,0,1,0),objective)$par

[1]  1.3154087 -0.4478625  0.3857039  1.3727617
```

(helper functions are needed because `optim()` optimizes over $\mathbb{R}^n$). This shows reasonable agreement with the true value of the mean and indeed the analytic value of the MLE, specifically

```
> colMeans(z)

[1] 1.315426-0.447472i 0.386068+1.372784i
```

3. The Gaussian process

In the context of the emulator, a (real) Gaussian process is usually defined in terms of a random function $\eta\colon \mathbb{R}^p \longrightarrow \mathbb{R}$ which, for any set of points $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ in its domain the random vector $\{\eta(\mathbf{x}_1), \ldots, \eta(\mathbf{x}_n)\}$ is multivariate Gaussian.

It is convenient to define means and variances as $\mathbb{E}\eta(\mathbf{x})|\,\boldsymbol{\beta} = h(\mathbf{x})\,\boldsymbol{\beta}$, conditional on the (unknown) vector of coefficients $\boldsymbol{\beta}$ and $h(\cdot)$, the q known regressor functions of $\mathbf{x} = (x_1, \ldots, x_p)^T$; a common choice is $h(\mathbf{x}) = (1, x_1, \ldots, x_p)^T$, but one is free to choose any function of $\mathbf{x}$. The covariance is typically given by

$$\mathsf{COV}\left(\eta(\mathbf{x}), \eta(\mathbf{x}')\right) = V\left(\mathbf{x} - \mathbf{x}'\right)$$

where $V: \mathbb{R}^n \longrightarrow \mathbb{R}$ must be chosen so that the variance matrix of any finite set of observations is always positive-definite. Bochner's theorem (Feller 1971, chapter XIX) shows that $V(\cdot)$ must be proportional to the characteristic function of a symmetric probability Borel measure.

3.1. Complex Gaussian processes

The complex case is directly analogous, with $\eta\colon \mathbb{C}^p \longrightarrow \mathbb{C}$ and $\boldsymbol{\beta} \in \mathbb{C}^q$. Writing $\mathsf{COV}(\eta(\mathbf{z}_1), \ldots, \eta(\mathbf{z}_n)) = \Omega$, so that element $(i,j)^{\text{th}}$ of matrix Ω is $\mathsf{COV}(\eta(\mathbf{z}_i), \ldots, \eta(\mathbf{z}_j))$, we may relax the reqirement that Ω be symmetric positive definite to requiring only Hermitian positive definiteness. This allows one to use the characteristic function of *any*, possibly non-symmetric, random variable with support over $\mathbb{C}^p$. That Ω remains Hermitian positive definite may be shown by evaluating a quadratic form with it and $\mathbf{w} \in \mathbb{C}^n$ and establishing that it is real and non-negative:

$$
\begin{aligned}
\mathbf{w}^{\star}\Omega\mathbf{w} &= \sum_{i,j}\overline{\mathbf{w}_i}\mathsf{COV}\left(\eta\left(\mathbf{x}_i\right),\eta\left(\mathbf{x}_j\right)\right)\mathbf{w}_j\\
&= \sum_{i,j}\overline{\mathbf{w}_i}\left[\int_{\mathbf{t}\in\mathbb{C}^n} e^{\mathrm{i}\,\mathrm{Re}\,\mathbf{t}^{\star}(\mathbf{x}_i-\mathbf{x}_j)}\,\Pr\left(\mathbf{t}\right)\,d\mathbf{t}\right]\mathbf{w}_j\\
&= \int_{\mathbf{t}\in\mathbb{C}^n}\left[\sum_{i,j}\overline{\mathbf{w}_i}e^{\mathrm{i}\,\mathrm{Re}\,\mathbf{t}^{\star}(\mathbf{x}_i-\mathbf{x}_j)}\mathbf{w}_j\,\Pr\left(\mathbf{t}\right)\right]\,d\mathbf{t}\\
&= \int_{\mathbf{t}\in\mathbb{C}^n}\left[\sum_{i,j}\overline{\mathbf{w}_i}e^{\mathrm{i}\,\mathrm{Re}(\mathbf{t}^{\star}\mathbf{x}_i)}\overline{\overline{\mathbf{w}_j}e^{\mathrm{i}\,\mathrm{Re}(\mathbf{t}^{\star}\mathbf{x}_j)}}\,\Pr\left(\mathbf{t}\right)\right]\,d\mathbf{t}\\
&= \int_{\mathbf{t}\in\mathbb{C}^n}\left|\sum_{i}\overline{\mathbf{w}_i}e^{\mathrm{i}\,\mathrm{Re}(\mathbf{t}^{\star}\mathbf{x}_i)}\right|^2\,\Pr\left(\mathbf{t}\right)\,d\mathbf{t}\\
&\geqslant 0.
\end{aligned}
$$

This motivates the characteristic function of a complex multivariate random variable $\mathbf{Z}$ is defined as $\mathbb{E}e^{i\,\mathrm{Re}(\mathbf{t}^{\star}\mathbf{Z})}$. Thus the covariance matrix is Hermitian positive definite: although its entries are not necessarily real, its eigenvalues are all nonnegative.

In the real case one typically chooses $V\left(\cdot\right)$ to be the characteristic function of a Gaussian distribution; in the complex case one can use the complex multivariate distribution 2 which has characteristic function

$$\exp\left(i\,\mathrm{Re}\left(\mathbf{t}^{\star}\boldsymbol{\mu}\right)-\frac{1}{4}\mathbf{t}^{\star}\Gamma\mathbf{t}\right) \tag{3}$$

and following Hankin (2012) in writing $\mathfrak{B}=\Gamma/4$, we can write the variance matrix as a product of a (real) scalar σ^2 term and

$$c\left(\mathbf{t}\right)=\exp\left(i\,\mathrm{Re}\left(\mathbf{t}^{\star}\boldsymbol{\mu}\right)-\mathbf{t}^{\star}\mathfrak{B}\mathbf{t}\right). \tag{4}$$

In equation 4, $\mathfrak{B}$ has the same meaning as in conventional emulator techniques and controls the modulus of the covariance between $\eta\left(\mathbf{z}\right)$ and $\eta\left(\mathbf{z}'\right)$; $\boldsymbol{\mu}$ governs the phase.

Given the above, it seems to be reasonable to follow Oakley (1999) and admit only diagonal $\mathfrak{B}$; but now distributions with nonzero mean can be considered. Such a parametrization gives $3p$ (real) hyperparameters; compare $2p$ if $\mathbb{C}^p$ is identified with $\mathbb{R}^{2p}$.

4. Functions of several complex variables

Analytic functions of several complex variables are an important and interesting class of objects; Krantz (1987) motivates and discusses the discipline. Formally, consider $f\colon\mathbb{C}^n\longrightarrow\mathbb{C}$, $n\geqslant 2$ and write $f\left(z_1,\ldots,z_n\right)$. Function f is *analytic* if it satisfies the Cauchy-Riemann conditions in each variable separately, that is $\partial f/\partial\overline{z}_j=0$, $1\leqslant j\leqslant n$.

Such an f is continuous (due to a "non-trivial theorem of Hartogs") and continuously differentiable to arbitrarily high order. Krantz goes on to state some results which are startling if one's exposure to complex analysis is restricted to functions of a single variable: for example, any isolated singularity is removable.

5. Numerical illustration of these ideas

The natural definition of complex Gaussian process above, together with the features of analytic functions of several complex variables, suggests that a complex emulation of analytic functions of several complex variables might be a useful technique.

The ideas presented above, and the **cmvnorm** package, can now be used to sample directly from an appropriate complex gaussian distribution and estimate the roughness parameters:

```
> val <- latin.hypercube(40,2,names=c('a','b'),complex = TRUE)
> head(val)

                    a                 b
[1,] 0.7375+0.2375i 0.2375+0.7125i
[2,] 0.6875+0.5875i 0.1375+0.3375i
[3,] 0.4625+0.5375i 0.9875+0.5875i
[4,] 0.7875+0.0625i 0.0625+0.7875i
[5,] 0.3875+0.0375i 0.5875+0.7625i
[6,] 0.2125+0.5625i 0.7625+0.9625i
```

and now specify a variance matrix using simple values for the roughness hyperparameters $\mathfrak{B} = \left(\begin{smallmatrix}1 & 0\\ 0 & 2\end{smallmatrix}\right)$ and $\boldsymbol{\mu} = (1, i)^T$:

```
> true_scales <- c(1,2)
> true_means <- c(1,1i)
> A <- corr_complex(val, means=true_means, scales=true_scales)
> round(A[1:4,1:4],2)

             [,1]        [,2]        [,3]        [,4]
[1,] 1.00+0.00i 0.59-0.27i 0.25-0.10i 0.89+0.11i
[2,] 0.59+0.27i 1.00+0.00i 0.20+0.00i 0.42+0.26i
[3,] 0.25+0.10i 0.20+0.00i 1.00+0.00i 0.10+0.06i
[4,] 0.89-0.11i 0.42-0.26i 0.10-0.06i 1.00+0.00i
```

Function `corr_complex()` is a complex generalization of `corr()`; matrix `A` is Hermitian positive-definite. It is now possible to make a single multivariate observation d of this process, using $\boldsymbol{\beta} = (1, 1+i, 1-2i)^T$:

```
> true_beta <- c(1,1+1i,1-2i)
> d <- drop(rcmvnorm(n=1,mean=regressor.multi(val) %*% true_beta,sigma=A))
> head(d)

[1] 3.212719+1.594901i 1.874278+0.345517i 3.008503-0.767618i 3.766526+2.071882i
[5] 3.712913+0.800983i 3.944167+0.924833i
```

thus `d` is a single observation from a complex multivariate Gaussian distribution. Most of the functions of the **emulator** package operate without modification:

```
> betahat.fun(val,solve(A),d)

               const                  a                  b
0.593632-0.0128655i 0.843608+1.0920437i 1.140372-2.5053751i
```

but the `interpolant()` functionality is implemented in the **cmvnorm** package (the likelihood function is different). So for example it is possible to evaluate the posterior distribution of the process at $(0.5, 0.3 + 0.1i)$, a point at which no observation has been made:

```
> interpolant.quick.complex(rbind(c(0.5,0.3+0.1i)),d,
+     val,solve(A),scales=true_scales,means=true_means,give.Z=TRUE)

$mstar.star
[1] 1.706402-1.008601i

$Z
[1] 0.203295

$prior
[1] 1.608085-0.104419i
```

Thus the posterior distribution for the process at this point is Gaussian with a mean of about $1.73 + 1.03i$ and a variance of about 0.16.

5.1. Analytic functions

These techniques are now used to emulate an analytic function of several complex variables. A complex function's being analytic is a very strong restriction; Needham (2004) uses 'rigidity' to describe the severe constraint that analyticity represents.

Here the Weierstrass sigma function is chosen as an example, on the grounds that Littlewood considers the σ-function to be a "typical" entire function. The elliptic package (Hankin 2006) is used for numerical evaluation. The σ-function takes a primary argument z and two invariants g_1, g_2, so a three-column complex design matrix is required:

```
> require("emulator")
> require("elliptic")
> valsigma <-
+     2+1i + round(latin.hypercube(30,3,names=c("z","g1","g2"),complex=TRUE)/4,2)
> head(valsigma)

              z         g1         g2
[1,] 2.17+1.15i 2.09+1.22i 2.21+1.09i
[2,] 2.11+1.01i 2.04+1.03i 2.25+1.15i
```

```
[3,] 2.10+1.04i 2.15+1.00i 2.22+1.20i
[4,] 2.13+1.10i 2.24+1.21i 2.01+1.16i
[5,] 2.20+1.00i 2.20+1.06i 2.08+1.08i
[6,] 2.05+1.10i 2.19+1.04i 2.11+1.03i
```

An offset to `latin.hypercube()` is needed because $\sigma(z, g_1, g_2) = z + \mathcal{O}(z^5)$. The sigma function can now be evaluated at the points on the design matrix:

```
> dsigma <- apply(valsigma,1,function(u){sigma(u[1],g=u[2:3])})
```

Function `scales.likelihood.complex()` can be used to return the log-likelihood for a specific set of roughness parameters:

```
> scales.likelihood.complex(scales=c(1,1,2),means=c(1,1+1i,1-2i),
+                           zold=valsigma,z=dsigma,give_log=TRUE)

[1] 144.5415
```

Numerical methods can then be used to find the maximum likelihood estimate. Because function `optim()` optimizes over $\mathbb{R}^n$, helper functions are again needed which translate from the optimand to scales and means:

```
> scales <- function(x){exp(x[c(1,2,2)])}
> means <-  function(x){x[c(3,4,4)] + 1i*x[c(5,6,6)]}
```

Because the diagonal elements of $\mathfrak{B}$ are strictly positive, their *logarithms* are optimized, following Hankin (2005); it is implicitly assumed that the scales and means associated with g_1 and g_2 are equal.

```
> objective <- function(x){
+  -scales.likelihood.complex(scales=scales(x),means=means(x),zold=valsigma,z=dsigma)
+ }
> start <-
+     c(-0.538, -5.668, 0.6633, -0.0084, -1.73, -0.028)
> jj <- optim(start,objective,method="SANN",control=list(maxit=100))
> (u <- jj$par)

[1] -0.5380 -5.6680  0.6633 -0.0084 -1.7300 -0.0280

> Asigma <- corr_complex(z1=valsigma,scales=scales(u),means=means(u))
```

So now we can compare the emulator against the "true" value:

```
> interpolant.quick.complex(rbind(c(2+1i,2+1i,2+1i)), zold=valsigma,
+     d=dsigma,Ainv=solve(Asigma),scales=scales(u),means=means(u))
```

```
[1] 3.078956+1.259993i
```

```
> sigma(2+1i,g=c(2+1i,2+1i))
```

```
[1] 3.078255+1.257819i
```

showing reasonable agreement. It is also possible to test the hypothesis $H_{\mathbb{R}}$ that the variance matrix A is real, by calculating the likelihood ratio of the full model 4 and that obtained by $H_{\mathbb{R}}$, that is, performing the optimization constrained so that $\boldsymbol{\mu} \in \mathbb{R}^2$:

```
> ob2 <- function(x){
+     -scales.likelihood.complex(scales=scales(x),means=c(0,0,0),zold=valsigma,z=dsigma)
+ }
> jjr <- optim(u[1:2],ob2,method="SANN",control=list(maxit=1000))
> (ur <- jjr$par)
```

```
[1]  0.2136577 -4.2640825
```

so the test statistic D is given by

```
> LR <- scales.likelihood.complex(scales=scales(ur),means=c(0,0,0),zold=valsigma,z=dsigma)
> LC <- scales.likelihood.complex(scales=scales(u),means=means(u),zold=valsigma,z=dsigma)
> (D <- 2*(LC-LR))
```

```
[1] 22.17611
```

Observing that D is in the tail region of its asymptotic distribution, χ^2_3, the hypothesis $H_{\mathbb{R}}$ may be rejected.

6. Conclusions

The **cmvnorm** package for the complex multivariate Gaussian distribution has been introduced and motivated. The Gaussian proces has been generalized to the complex case, and a complex generalization of the emulator technique has been applied to an analytic function of several complex variables. The complex variance matrix was specified using a novel parameterization which accommodated non-real covariances in the context of circulary symmetric random variables. Further work might include numerical support for the complex multivariate Student t distribution.

References

Feller W (1971). *An introduction to probability theory and its applications*, volume 2. Wiley.

Genz A, Bretz F, Miwa T, Mi X, Leisch F, Scheipl F, Hothorn T (2014). ***mvtnorm**: Multivariate Normal and t Distributions*. R package version 1.0-0, URL http://CRAN.R-project.org/package=mvtnorm.

Goodman NR (1963). "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)." *The Annals of Mathematical Statistics*, **34**(1), 152–177.

Hankin RKS (2005). "Introducing **BACCO**, an R bundle for Bayesian analysis of computer code output." *Journal of Statistical Software*, **14**(16).

Hankin RKS (2006). "Introducing **elliptic**, an R package for elliptic and modular functions." *Journal of Statistical Software*, **15**(7).

Hankin RKS (2012). "Introducing **multivator**: a multivariate emulator." *Journal of Statistical Software*, **46**(8).

Krantz SG (1987). "What is several complex variables?" *The American Mathematical Monthly*, **94**(3), 236–256.

Needham T (2004). *Visual complex analysis*. Clarendon Press, Oxford.

Oakley J (1999). *Bayesian uncertainty analysis for complex computer codes*. Ph.D. thesis, University of Sheffield.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Affiliation:

Robin K. S. Hankin
Auckland University of Technology
2-14 Wakefield Street
Auckland NZ
hankin.robin@gmail.com