NONLINEAR STATISTICAL SIGNAL PROCESSING: USEFUL THEOREMS AND THEIR APPLICATION

A. van den Bos

Delft University of Technology, Department of Applied Physics
P.O. Box 5046, 2600 GA Delft, The Netherlands
e-mail: a.vandenbos@tn.tudelft.nl

ABSTRACT

A coherent treatment is presented of Price's theorem, Pawula's theorem and Bonnet's theorem. These theorems are often used for the computation of the expectation of nonlinear functions of normally distributed stochastic variables or stochastic processes. In addition, a number of modifications and simplifications of these theorems is proposed. Finally, a class of non-normal distributions for which the theorems are true is briefly discussed.

1. INTRODUCTION

Price's theorem [1], Pawula's theorem [2] and, to a lesser extent, Bonnet's theorem [3] are proven tools in nonlinear statistical signal processing. These theorems are all concerned with the computation of the expectation of nonlinear functions of normally distributed variates. These expectations are needed in many applications. This is reflected by the literature where applications of the theorems are reported in the study of receivers, correlation techniques, spectral analysis, amplitude quantization, phase locked loop design and analysis, adaptive ltering and perceptron studies. The theorems often produce closed form expressions or expressions suitable for numerical evaluation, even for expectations of complicated multivariable functions.

Price's theorem relates the expectation of a nonlinear function of a number of normally distributed variates to the covariances of the variates. This relation has the form of a set of partial differential equations with the covariances as independent variables. The expectation of the nonlinear function is the solution of the set. The theorem also has a complex valued version [4].

Pawula's theorem is an alternative to Price's theorem. Here, the covariance matrix of the variates is parameterized by a single scalar parameter. Instead of a set of partial differential equations, an ordinary linear differential equation with the parameter as independent variable must be solved for the expectation of the nonlinear function.

Finally, Bonnet's theorem describes the relation of the expectation of a nonlinear function of normal variates to their expectations.

The three theorems are not, or not coherently, treated in the reference books known to the author. Therefore, a purpose of the paper is to briefly review them and to present them in a unified way. Furthermore, a number of modifications and simplifications is proposed. For example, the impractical assumption in the usual presentation of Price's and Pawula's theorem that the variances of the variates are all equal to one is removed. Furthermore, Price's theorem is extended to include variances of the variates instead of covariances only. Also, Pawula's theorem is modified to simplify the generation of the integration constant concerned.

Finally, the applicability of the theorems to sums of normal variates and independent non-normal ones is briefly discussed. As an example, the proof that Price's theorem is true for such sums is outlined.

2. THREE THEOREMS

In this section, Price's theorem, Pawula's theorem and Bonnet's theorem will be described. In these descriptions, x_1, x_2 $,..., x_N$ are jointly normally distributed variates with m_n the expectation of x_n and c_{pq} the covariance of x_p and x_q. Furthermore, f will denote a, nonlinear, scalar function of $x_1, x_2, ..., x_N$. Then *Price's theorem* [1] states that

$$\frac{\partial E[f]}{\partial \rho_{pq}} = E[\frac{\partial^2 f}{\partial x_p \partial x_q}] \qquad (1)$$

where $p \neq q = 1, ..., N$ and the expectations are taken with respect to the multivariable normal distribution with all variances equal to one and with correlation coefficients ρ_{pq}. Higher-order derivatives with respect to any of the ρ_{pq} follow from (1) by recursion. Equation (1) is a first-order partial differential which has to be solved if the purpose is the computation of the expectation of the function.

Example. The arcsine relation.

Suppose that

$$f = sgn(x_1)sgn(x_2)$$

where the sign function $sgn(.)$ is equal to 1, 0, or -1 as its argument is positive, equal to zero, or negative, respectively. Also suppose that m_1 or m_2 is equal to zero. Then

$$\begin{aligned} \frac{\partial E[f]}{\partial \rho_{12}} &= 4E[\delta(x_1)\delta(x_2)] \\ &= \frac{2}{\pi}\frac{1}{(1-\rho_{12}^2)^{\frac{1}{2}}} \end{aligned}$$

where $\delta(.)$ is the Dirac delta function. Hence

$$E[f] = \frac{2}{\pi}\arcsin \rho_{12} + constant.$$

The integration constant is equal to zero since for $\rho_{12} = 0$

$$E[f] = E[sgn(x_1)]\ E[sgn(x_2)] = 0$$

since m_1 or m_2 is equal to zero. □

This example shows an important characteristic of Price's theorem for functions of two variables: relatively complicated integrations required by direct computation are avoided and replaced by a, sometimes simple, solution of a first order partial differential equation in the covariance of the variates.

With respect to Price's theorem, the following remarks may be made. First, the variances are supposed to be equal to one which is impractical. Also, as a consequence, derivatives with respect to the variances are not included.

If f is a function of more than two variates, the computation of its expectation requires, in principle, the solution of a *set* of partial differential equations. In *Pawula's theorem* [2], on the other hand, the covariance matrix of the variates is parameterized by a single scalar parameter. The computation of the expectation of the function requires the solution of a single ordinary differential equation in this parameter for any number of variates. The parameterization of the covariance matrix is as follows. The original covariance matrix $\rho = [\rho_{pq}]$ with $\rho_{pp} = 1$, $p, q = 1, ..., N$ is replaced by

$$\rho_\alpha = [\alpha^{1-\delta_{pq}}\rho_{pq}]. \tag{2}$$

with δ_{pq} is the Kronecker delta. This implies that the diagonal elements of the new covariance matrix ρ_α are equal to one and the remaining elements are equal to $\alpha\rho_{pq}$, respectively. Then Pawula's theorem states that

$$\frac{dE_\alpha[f]}{d\alpha} = \sum_{p<q} \rho_{pq}\ E_\alpha[\frac{\partial^2 f}{\partial x_p \partial x_q}]\ .$$

As a solution

$$E[f] = E_0[f] + \sum_{p<q} \rho_{pq} \int_0^1 E_\alpha[\frac{\partial^2 f}{\partial x_p \partial x_q}]d\alpha$$

is taken where $E_0[f]$ is the expectation of the function f if all x_p are uncorrelated. $E_0[f]$ is not necessarily easy to compute. From this description, it is clear that again the impractical assumption is made that all variances are equal to one.

Both Price's and Pawula's theorem are concerned with parameterizations of the covariance matrix. In Price's theorem, the parameters are the off-diagonal covariances; in Pawula's theorem the parameter is the multiplier α. In *Bonnet's theorem* [3], on the other hand , the expectations $m_1, ...,$ m_N of the variates $x_1, ..., x_N$ are taken as parameters. Bonnet's theorem states that

$$\frac{\partial E[f]}{\partial m_r} = E[\frac{\partial f}{\partial x_r}]$$

with $r = 1, ..., N$.

The proofs of the three theorems in the original references are highly diverse and that of Price's theorem is relatively complicated. In any case, the theorems have in common that they produce a derivative of the desired function expectation with respect to one or more parameters of the normal probability density function. Using the fact that a parameter of the probability density function is also a parameter of the characteristic function, Papoulis [5] has considerably simplified the proof of Price's theorem. This proof concerns functions of two normally distributed variates and covariances only. In this paper, Papoulis's idea is extended to include any number of variates and variances as well as covariances. It is also used to outline a proof that Price's theorem is also true for sums of normal and independent non-normal variates.

3. PROOFS USING CHARACTERISTIC FUNCTIONS

Define x as the vector of normal variates

$$x = (x_1...x_N)^T$$

and

$$m = (m_1...m_N)^T$$

as the vector of their expectations, respectively, where the superscript T indicates transposition. Furthermore, define the covariance matrix of x as $C = [c_{pq}]$ where c_{pq} is the covariance of x_p and x_q.Then the probability density function of x is described by

$$p(x) = \frac{1}{(2\pi)^{\frac{N}{2}}(\det C)^{\frac{1}{2}}}\exp\{-\frac{1}{2}(x-m)^T C^{-1}(x-m)\}.$$

The corresponding characteristic function $P(\omega)$ is the conjugate of the Fourier transform of $p(x)$ [5]

$$P(\omega) = \exp(-\frac{1}{2}\omega^T C\omega + j\omega^T m) \tag{3}$$

where $j^2 = -1$ and $\omega = (\omega_1...\omega_N)^T$ is the vector of independent variables of $P(\omega)$. Therefore,

$$p(x) = \frac{1}{(2\pi)^N} \int P(\omega) \exp(-j\omega^T x) d\omega \qquad (4)$$

where the integration is over $\omega_1, ..., \omega_N$.

3.1. A proof of Price's theorem

A modified version of Price's theorem may be derived as follows. By (4)

$$\begin{aligned} E[f] &= \int f(x)p(x)dx & (5) \\ &= \int f(x) \frac{1}{(2\pi)^N} \int P(\omega) \exp(-j\omega^T x) d\omega \, dx \end{aligned}$$

and hence

$$\frac{\partial E[f]}{\partial c_{pq}} = \int f(x) \frac{1}{(2\pi)^N} \int \frac{\partial P(\omega)}{\partial c_{pq}} \exp(-j\omega^T x) d\omega \, dx \,. \qquad (6)$$

From (3) it follows that

$$\frac{\partial P(\omega)}{\partial c_{pq}} = -(\frac{1}{2})^{\delta_{pq}} \omega_p \omega_q P(\omega).$$

Substituting of this result in (6) and, subsequently, applying the Fourier differentiation theorem shows that

$$\frac{\partial E[f]}{\partial c_{pq}} = (\frac{1}{2})^{\delta_{pq}} E[\frac{\partial^2 f}{\partial x_p \partial x_q}]. \qquad (7)$$

This version of Price's theorem is different from (1) in two respects: the variances need not be equal to one and the partial derivatives of $E[f]$ with respect to the variances c_{pp} are included. This enhances the applicability of the theorem.

3.2. A proof of Pawula's theorem

In this section, a proof of Pawula's theorem using a parameterization different from that of (2) is given. The parameterization proposed is:

$$C_\alpha = \alpha[c_{pq}]. \qquad (8)$$

If in (2) α is equal to zero, the covariance matrix of x becomes the identity matrix. If, on the other hand, in (8) α is equal to zero, the covariance matrix of x is the null matrix. Then the corresponding probability density function is singular and becomes a Dirac delta function located at the point $(m_1, ..., m_N)$. Furthermore, the variances c_{pp} in (8) may be different from one. The characteristic function of normally distributed variates with a covariance matrix (8) is described by

$$P_\alpha(\omega) = \exp(-\frac{1}{2}\alpha\omega^T C\omega + j\omega^T m). \qquad (9)$$

Then

$$\frac{\partial E_\alpha[f]}{\partial \alpha} = \int f(x) \frac{1}{(2\pi)^N} \int \frac{\partial P_\alpha(\omega)}{\partial \alpha} \exp(-j\omega^T x) d\omega \, dx \,. \qquad (10)$$

From (9) it follows that

$$\frac{\partial P_\alpha(\omega)}{\partial \alpha} = -\frac{1}{2}\omega^T C\omega \, P_\alpha(\omega).$$

Substituting this result in (10) and applying the Fourier differentiation theorem yields

$$\frac{dE_\alpha[f]}{d\alpha} = \frac{1}{2} \int f(x) \sum_{p,q} c_{pq} \frac{\partial^2 p_\alpha(x)}{\partial x_p \partial x_q} dx$$

Finally, integrating by parts over x_p and x_q shows that

$$\frac{dE_\alpha[f]}{d\alpha} = \frac{1}{2} \sum_{p,q} c_{pq} E_\alpha[\frac{\partial^2 f}{\partial x_p \partial x_q}].$$

Integration with respect to α then produces

$$E[f] = f(m) + \frac{1}{2} \int_0^1 \sum_{p,q} c_{pq} E_\alpha[\frac{\partial^2 f}{\partial x_p \partial x_q}] d\alpha.$$

Notice that the integration constant $f(m)$ only requires the substitution $m = (m_1...m_N)^T$ for x in $f(x)$.

3.3. A proof of Bonnet's theorem

Bonnet's theorem may be derived as follows. From (5) it follows that

$$\frac{\partial E[f]}{\partial m_r} = \int f(x) \frac{1}{(2\pi)^N} \int \frac{\partial P(\omega)}{\partial m_r} \exp(-j\omega^T x) d\omega \, dx \qquad (11)$$

with $r = 1, ..., N$. Furthermore, (3) shows that

$$\frac{\partial P(\omega)}{\partial m_r} = j\omega_r P(\omega).$$

Then substituting this expression in (11) and applying the Fourier differentiation theorem shows that

$$\frac{\partial E[f]}{\partial m_r} = \int f(x) \frac{\partial p(x)}{\partial x_r} dx.$$

Integrating by parts over x_r yields Bonnet's theorem:

$$\frac{\partial E[f]}{\partial m_r} = E[\frac{\partial f}{\partial x_r}] \,.$$

3.4. General characteristics of the proofs

The three proofs presented in the sections 3.1 -3.3 are based on suitably parameterized characteristic functions. In Price's theorem, the parameters are all variances and covariances of the normal variates concerned. In Pawula's theorem, there is only one scalar parameter by which the covariance matrix is multiplied while in Bonnet's theorem the parameters are all expectations of the variates. Other parameterizations, such as those used in Price's and Pawula's original publications, are also possible. Then in each of the proofs the probability density function is written as the inverse Fourier transform of the characteristic function. The remaining steps are the same in each of the theorems and are standard calculus. Since thus the only difference in the three theorems is the parameterization of the characteristic function, it is believed that a considerable unification and simplification has been achieved. In addition, in the proofs impractical assumptions such as variances equal to one were found to be unnecessary.

4. NON-NORMAL VARIATES

In a comment on his own paper [6], Price states without proof that his theorem is also true for sums of normally distributed variates and independent non-normal variates if the derivatives concerned are taken with respect to the covariances of the normal variates. Analogously, Pawula's theorem and Bonnet's theorem may be extended to include such sum variates. As a representative example, the proof of Price's theorem for sum variates is presented now.

Suppose that independent, non-normal variates $y = (y_1 ...y_N)^T$ with a characteristic function $Q(\omega)$ are added to the normally distributed variates $x = (x_1...x_N)^T$. Then the characteristic function $R(\omega)$ of the of the sum variates $z = x + y$ is described by

$$R(\omega) = \exp(-\frac{1}{2}\omega^T C\omega + j\omega^T m)\ Q(\omega).$$

From this expression, it follows that

$$\frac{\partial R(\omega)}{\partial c_{pq}} = -(\frac{1}{2})^{\delta_{pq}}\omega_p\omega_q R(\omega).$$

The rest of the proof is the same as the proof of Price's theorem given in section 3.1. The result is:

$$\frac{\partial E[f]}{\partial c_{pq}} = (\frac{1}{2})^{\delta_{pq}} E[\frac{\partial^2 f}{\partial z_p \partial z_q}]$$

where the expectations are taken with respect to z .

5. CONCLUSIONS

It has been shown that Price's theorem, Pawula's theorem and Bonnet's theorem can be proved and formulated in a unified and simplified way. Also the existence of non-normally distributed variates to which the theorems may be applied has been briefly discussed.

6. REFERENCES

[1] R. Price. A useful theorem for nonlinear devices having Gaussian inputs. *IRE Transactions on Information Theory*, IT-4(2):69–72, June 1958.

[2] R.F. Pawula. A modified version of Price's theorem. *IEEE Transactions on Information Theory*, IT-13(2):285–288, April 1967.

[3] G. Bonnet. Sur certaines propriétés statistiques de fonctions aléatoires issues de transformations non linéaires. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 258:4917–4920, May 1964.

[4] A. van den Bos. Price's theorem for complex variates. *IEEE Transactions on Information Theory*, 42(1):286–287, January 1996.

[5] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.

[6] R. Price. Comment on: ' A useful theorem for nonlinear devices having Gaussian inputs'. *IEEE Transactions on Information Theory*, IT-10(2):171, April 1964.