LINEAR TRANSFORMATIONS & THE MULTIVARIATE GENERATING FUNCTION

MICHAEL C. BURKHART

Abstract. This note examines linear combinations of multi-indexed sequences and derives the multivariate generating function of such a linear combination in terms of the original sequence's m.g.f. Applications include finding distributions and moments of non-negative discrete random variables conditioned on non-negative linear combinations of the original variables. Examples include independent Poisson r.v.'s and a d-variate multinomial distribution.

1. Introduction

Set $\mathbb{N} = \mathbb{Z}_{\geq 0}$. Let $\mathfrak{b} : \mathbb{N}^d \to \mathbb{C}$ denote a multi-indexed sequence of complex numbers. Then the multivariate generating function for $\mathfrak{b}$ is given by:

$$G_{\mathfrak{b}}(t_1, \ldots, t_d) = \sum_{(j_1,\ldots,j_d)\in\mathbb{N}^d} \mathfrak{b}_{(j_1,\ldots,j_d)} t_1^{j_1} \cdots t_d^{j_d}$$

Such generating functions have applications throughout discrete mathematics, especially to combinatorial classes and probability distributions [FS09], and prove useful in finding recurrences, moments, and asymptotics [Wil94, PW08]. Multivariate generating functions can be used to obtain conditional distributions [Xek87, JKB96]. Applications to biochemistry include stochastic models of chemical network theory, in particular chemical kinetics [SZ10].

2. Main Result

Fix some matrix $\mathbf{A} = (a_{ij}) \in \mathbf{Mat}_{m\times d}(\mathbb{N})$. Define a new multi-indexed sequence $\mathfrak{c} : \mathbb{N}^m \to \mathbb{C}$ by taking linear combinations of the sequence $\mathfrak{b}$ using the coefficients of the matrix $\mathbf{A}$. That is, for $(k_1, \ldots, k_m) \in \mathbb{N}^m$ set:

$$\mathfrak{c}_{(k_1,\ldots,k_m)} = \sum_{\substack{(j_1,\ldots,j_d)\in\mathbb{N}^d \\ (k_1,\ldots,k_m)^T = \mathbf{A}(j_1,\ldots,j_d)^T}} \mathfrak{b}_{(j_1,\ldots,j_d)}$$

Theorem 1. *The analogously-defined multivariate generating function for the sequence $\mathfrak{c}$ is given by:*

$$G_{\mathfrak{c}}(z_1, \ldots, z_m) = G_{\mathfrak{b}}(\Pi_{i=1}^m z_i^{a_{i1}}, \ldots, \Pi_{i=1}^m z_i^{a_{id}})$$

where $\mathbf{A} = (a_{ij})$.

2010 *Mathematics Subject Classification.* 05A15, 32A05, 60–08.

Proof. This proof mirrors Sontag and Zeilberger's proof of the special case where $\mathfrak{b}_{(j_1,\dots,j_d)}$ is the joint probability distribution for independent Poisson random variables [SZ10]. From the definition:

$$
\begin{aligned}
G_{\mathfrak{c}}(z_1,\dots,z_m) &= \sum_{(k_1,\dots,k_m)\in\mathbb{N}^m} \mathfrak{c}_{(k_1,\dots,k_m)} z_1^{k_1}\cdots z_m^{k_m} \\
&= \sum_{(k_1,\dots,k_m)\in\mathbb{N}^m} \left(\sum_{\substack{(j_1,\dots,j_d)\in\mathbb{N}^d \\ (k_1,\dots,k_m)^T=\mathbf{A}(j_1,\dots,j_d)^T}} \mathfrak{b}_{(j_1,\dots,j_d)} \right) z_1^{k_1}\cdots z_m^{k_m}
\end{aligned}
$$

Exchaning the order of summation then yields:

$$
\begin{aligned}
G_{\mathfrak{c}}(z_1,\dots,z_m) &= \sum_{(j_1,\dots,j_d)\in\mathbb{N}^d} \left(\sum_{\substack{(k_1,\dots,k_m)\in\mathbb{N}^m \\ (k_1,\dots,k_m)^T=\mathbf{A}(j_1,\dots,j_d)^T}} \mathfrak{b}_{(j_1,\dots,j_d)} z_1^{k_1}\cdots z_m^{k_m} \right) \\
&= \sum_{(j_1,\dots,j_d)\in\mathbb{N}^d} \mathfrak{b}_{(j_1,\dots,j_d)} z_1^{a_{11}j_1+\cdots+a_{1d}j_d} \cdots z_m^{a_{m1}j_1+\cdots+a_{md}j_d} \\
&= \sum_{(j_1,\dots,j_d)\in\mathbb{N}^d} \mathfrak{b}_{(j_1,\dots,j_d)} (z_1^{a_{11}}\cdots z_m^{a_{m1}})^{j_1}\cdots(z_1^{a_{1d}}\cdots z_m^{a_{md}})^{j_d} \\
&= G_{\mathfrak{b}}(z_1^{a_{11}}\cdots z_m^{a_{m1}},\dots,z_1^{a_{1d}}\cdots z_m^{a_{md}}) = G_{\mathfrak{b}}(\Pi_{i=1}^m z_i^{a_{i1}},\dots,\Pi_{i=1}^m z_i^{a_{id}})
\end{aligned}
$$

□

3. Probability Generating Functions

Let $X_1,\dots,X_d$ be non-negative discrete random variables (not necessarily independent). The multivariate probability generating function (henceforth denoted p.g.f.) of $X_1,\dots,X_d$ is then given by:

$$
G_{\mathbf{X}}(t_1,\dots,t_d) = \sum_{(j_1,\dots,j_d)\in\mathbb{N}^d} \mathbb{P}(X_1=j_1,\dots,X_d=j_d)\; t_1^{j_1}\cdots t_d^{j_d}
$$

Define new random variables $Y_1,\dots,Y_m$ by taking linear combinations of the X_i:

$$
\begin{pmatrix} Y_1 \\ \vdots \\ Y_m \end{pmatrix} = \mathbf{A} \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} \tag{$\star$}
$$

Proposition 1. *The analogously-defined multivariate p.g.f. for $Y_1,\dots,Y_m$ is given by:*

$$
\begin{aligned}
G_{\mathbf{Y}}(z_1,\dots,z_m) &= \sum_{(k_1,\dots,k_m)\in\mathbb{N}^m} \mathbb{P}(Y_1=k_1,\dots,Y_m=k_m)\; z_1^{k_1}\cdots z_m^{k_m} \\
&= G_{\mathbf{X}}(\Pi_{i=1}^m z_i^{a_{i1}},\dots,\Pi_{i=1}^m z_i^{a_{id}})
\end{aligned}
$$

Apply the theorem to the multi-indexed sequences $\mathfrak{b}_{(j_1,\dots,j_d)} = \mathbb{P}(X_1=j_1,\dots,X_d=j_d)$ and $\mathfrak{c}_{(k_1,\dots,k_m)} = \mathbb{P}(Y_1=k_1,\dots,Y_m=k_m)$.

Independence. When in addition $X_1, \dots, X_d$ are independent non-negative discrete random variables, the p.g.f. takes the form:

$$G_{\mathbf{X}}(t_1, \dots, t_d) = \prod_{r=1}^{d} \Big(\sum_{j_r \in \mathbb{N}} \mathbb{P}(X_r = j_r)\ t_r^{j_r} \Big) = \prod_{r=1}^{d} G_{X_r}(t_r)$$

where G_{X_r} is the single-variable probability generating function for X_r. It follows that the p.g.f. for the linear combinations $Y_1, \dots, Y_m$ is given by:

$$G_{\mathbf{Y}}(z_1, \dots, z_m) = \prod_{r=1}^{d} \Big(G_{X_r}(\Pi_{i=1}^{m} z_i^{a_{ir}}) \Big)$$

Example. When $X_1, \dots, X_d$ are independent Poisson random variables (*cf.* [SZ10]), $X_r \sim Poisson(\lambda_r)$, $1 \le r \le n$:

$$G_{\mathbf{Y}}(z_1, \dots, z_m) = \exp \left(\sum_{r=1}^{d} \lambda_r \Big(\prod_{i=1}^{m} z_i^{a_{ir}} - 1 \Big) \right)$$

Simply note that, for $1 \le r \le n$:

$$G_{X_r}(t_r) = \sum_{i \ge 0} \frac{\lambda_r^i e^{-\lambda_r}}{i!} t_r^i = e^{-\lambda_r} \sum_{i \ge 0} \frac{(\lambda_r t_r)^i}{i!} = \exp(\lambda_r t_r - \lambda_r)$$

The Conditional Distribution of $\mathbf{X} \mid \mathbf{Y}$

Proposition 2. *Let the random vectors $\mathbf{X}, \mathbf{Y}$ be given as in $(\star)$. Then the joint multivariate p.g.f. for $\mathbf{X}, \mathbf{Y}$ is given:*

$$\begin{aligned} &G_{\mathbf{X},\mathbf{Y}}(t_1, \dots, t_d; z_1, \dots, z_m) \\ &\quad = \sum_{\substack{(j_1, \dots, j_d; k_1, \dots, k_m) \in \mathbb{N}^{d+m} \\ (k_1, \dots, k_m)^T = \mathbf{A}(j_1, \dots, j_d)^T}} \mathbb{P}(\mathbf{X} = \mathbf{j}, \mathbf{Y} = \mathbf{k})\ t_1^{d_1} \cdots t_d^{j_d} \cdot z_1^{k_1} \cdots z_m^{k_m} \\ &\quad = G_{\mathbf{X}}(t_1 \cdot \Pi_{i=1}^{m} z_i^{a_{i1}}, \dots, t_d \cdot \Pi_{i=1}^{m} z_i^{a_{id}}) \end{aligned}$$

This follows from noting that $\mathbb{P}(\mathbf{X} = \mathbf{j}, \mathbf{Y} = \mathbf{k}) = \mathbb{P}(\mathbf{X} = \mathbf{j})$ and then applying the theorem with $\mathfrak{b}_{(j_1, \dots, j_d)} = \mathbb{P}(\mathbf{X} = \mathbf{j})\ t_1^{d_1} \cdots t_d^{j_d}$. From joint multivariate p.g.f., it is possible to obtain the conditional p.g.f. of $\mathbf{X}$ given $Y_1 = k_1, \dots, Y_m = k_m$ [Xek87]. Further, the conditional pure (resp. mixed) factorial momements correspond to taking the coefficient of $z_1^{k_1} \cdots z_m^{k_m}$ in $G_{\mathbf{X},\mathbf{Y}}$ (which will be a polynomial in $t_1, \dots, t_d$), taking pure (resp. mixed) partial derivatives with respect to the t_r, evaluating at $t_1 = \cdots = t_d = 1$, and dividing by $\mathbb{P}(\mathbf{Y} = \mathbf{k})$. Let $[z_1^{j_1} \cdots z_d^{j_d}]G(z_1, \dots, z_d)$ denote the process of extracting the coefficient of $z_1^{j_1} \cdots z_d^{j_d}$ in the formal power series $G(z_1, \dots, z_d) = \sum_{(j_1, \dots, j_d) \in \mathbb{N}^d} G_{(j_1, \dots, j_d)} z_1^{j_1} \cdots z_d^{j_d}$. With this notation, it follows that:

$$\begin{aligned}
&\mathbb{E}(X_1^{(s_1)} \cdots X_d^{(s_d)} \mid Y_1 = k_1, \dots, Y_m = k_m) \\
&\qquad := \mathbb{E}\left(\frac{X_1!}{(X_1 - s_1)!} \cdots \frac{X_d!}{(X_d - s_d)!} \;\Big|\; Y_1 = k_1, \dots, Y_m = k_m \right) \\
&= \sum_{\substack{(j_1,\dots,j_d) \in \mathbb{N}^d \\ (k_1,\dots,k_m)^T = \mathbf{A}(j_1,\dots,j_d)^T}} \frac{j_1!}{(j_1 - s_1)!} \cdots \frac{j_d!}{(j_d - s_d)!} \; \mathbb{P}(X_1 = j_1, \dots, X_d = j_d \mid \mathbf{Y}) \\
&\qquad = \frac{[z_1^{k_1} \cdots z_m^{k_m}] \left(\frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} G_{\mathbf{X},\mathbf{Y}}(1, \dots, 1; z_1, \dots, z_m) \right)}{[z_1^{k_1} \cdots z_m^{k_m}] G_{\mathbf{Y}}(z_1, \dots, z_m)}
\end{aligned}$$

Combining this with Propositions 1 and 2 gives:

Theorem 2. *The conditional factorial moments of* $\mathbf{X}$ *given that* $Y_1 = k_1, \dots, Y_m = k_m$ *are:*

$$\begin{aligned}
&\mathbb{E}(X_1^{(s_1)} \cdots X_d^{(s_d)} \mid \mathbf{Y}) \\
&\quad = \frac{[z_1^{k_1} \cdots z_m^{k_m}] \left(\frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} G_{\mathbf{X}}(t_1 \cdot \Pi_{i=1}^m z_i^{a_{i1}}, \dots, t_d \cdot \Pi_{i=1}^m z_i^{a_{id}}) \right) \big|_{t_1 = \cdots = t_d = 1}}{[z_1^{k_1} \cdots z_m^{k_m}] G_{\mathbf{X}}(\Pi_{i=1}^m z_i^{a_{i1}}, \dots, \Pi_{i=1}^m z_i^{a_{id}})}
\end{aligned}$$

Example. If again $X_1, \dots, X_d$ are independent Poisson random variables (*cf.* [SZ10]), $X_r \sim Poisson(\lambda_r)$, $1 \le r \le n$, then:

$$\begin{aligned}
&\frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} G_{\mathbf{X}}(t_1 \cdot \Pi_{i=1}^m z_i^{a_{i1}}, \dots, t_d \cdot \Pi_{i=1}^m z_i^{a_{id}}) \\
&\qquad = \frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} \exp\left(\sum_{r=1}^d \lambda_r \Big(t_r \cdot \prod_{i=1}^m z_i^{a_{ir}} - 1 \Big) \right) \\
&\qquad = \left(\prod_{r=1}^d \left(\lambda_r \prod_{i=1}^m z_i^{a_{ir}} \right)^{s_r} \right) \exp\left(\sum_{r=1}^d \lambda_r \Big(t_r \cdot \prod_{i=1}^m z_i^{a_{ir}} - 1 \Big) \right)
\end{aligned}$$

So that:

$$\begin{aligned}
&[z_1^{k_1} \cdots z_m^{k_m}] \left(\frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} G_{\mathbf{X}}(t_1 \cdot \Pi_{i=1}^m z_i^{a_{i1}}, \dots, t_d \cdot \Pi_{i=1}^m z_i^{a_{id}}) \right) \big|_{t_1 = \cdots = t_d = 1} \\
&\quad = \prod_{r=1}^d \lambda_r^{s_r} \cdot [z_1^{k_1} \cdots z_m^{k_m}] \left\{ \prod_{i=1}^m z_i^{(\sum_{r=1}^d a_{ir} s_r)} \cdot \exp\Big(\sum_{r=1}^d \lambda_r \Big(\prod_{i=1}^m z_i^{a_{ir}} - 1 \Big) \Big) \right\} \\
&\quad = \prod_{r=1}^d \lambda_r^{s_r} \cdot [z_1^{k_1 - \sum_{r=1}^d a_{1r} s_r} \cdots z_m^{k_m - \sum_{r=1}^d a_{dr} s_r}] G_{\mathbf{Y}}(z_1, \dots, z_m)
\end{aligned}$$

Thus:

$$\begin{aligned}
&\mathbb{E}(X_1^{(s_1)} \cdots X_d^{(s_d)} \mid \mathbf{Y}) \\
&\qquad = \prod_{r=1}^d \lambda_r^{s_r} \cdot \frac{[z_1^{k_1 - \sum_{r=1}^d a_{1r} s_r} \cdots z_m^{k_m - \sum_{r=1}^d a_{dr} s_r}] G_{\mathbf{Y}}(z_1, \dots, z_m)}{[z_1^{k_1} \cdots z_m^{k_m}] G_{\mathbf{Y}}(z_1, \dots, z_m)}
\end{aligned}$$

whenever $k_i - \sum_{r=1}^{d} a_{ir}s_r \geq 0$ for all i and 0 otherwise. For computations on explicit matrices $\mathbf{A}$, Sontag and Zeilberger developed a Maple package utilizing Wilf-Zeilberger Theory to obtain moments and recurrences on the distribution [SZ10].

Example. When $X_1, \ldots, X_d$ have a d-variate multinomial distribution [JKB96, pp. 31-92] for some $N \in \mathbb{N}$ and $0 \leq p_1, \ldots, p_d \leq 1$ where $\sum_{i=1}^{d} p_i = 1$, the multivariate generating function is given:

$$G_{\mathbf{X}}(t_1, \ldots, t_d) = \sum_{\substack{(j_1,\ldots,j_d)\in\mathbb{N}^d \\ j_1+\cdots+j_d=N}} \binom{N}{j_1, \ldots, j_d} p_1^{j_1} \cdots p_d^{j_d} \cdot t_1^{j_1} \cdots t_d^{j_d}$$

$$= (p_1t_1 + \cdots + p_dt_d)^N$$

It follows that:

$$\frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} G_{\mathbf{X}}(t_1 \cdot \Pi_{i=1}^{m} z_i^{a_{i1}}, \ldots, t_d \cdot \Pi_{i=1}^{m} z_i^{a_{id}})$$

$$= \frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} (p_1t_1\Pi_{i=1}^{m} z_i^{a_{i1}} + \cdots + p_dt_d\Pi_{i=1}^{m} z_i^{a_{id}})^N$$

$$= \frac{N! \prod_{r=1}^{d} (p_r \prod_{i=1}^{m} z_i^{a_{ir}})^{s_r}}{(N - \sum_{r=1}^{d} s_r)!} \left(\sum_{r=1}^{d} p_rt_r\Pi_{i=1}^{m} z_i^{a_{ir}}\right)^{N-\sum_{r=1}^{d} s_r}$$

Whence:

$$(\dagger) \quad [z_1^{k_1} \cdots z_m^{k_m}] \left(\frac{\partial^{s_1}}{\partial t_1^{s_1}} \cdots \frac{\partial^{s_d}}{\partial t_d^{s_d}} G_{\mathbf{X}}(t_1 \cdot \Pi_{i=1}^{m} z_i^{a_{i1}}, \ldots, t_d \cdot \Pi_{i=1}^{m} z_i^{a_{id}})\right)\Big|_{t_1=\cdots=t_d=1}$$

$$= \frac{N! \prod_{r=1}^{d} p_r^{s_r}}{(N-\sum_{r=1}^{d} s_r)!} \cdot [z_1^{k_1} \cdots z_m^{k_m}] \Big\{ \left(\Pi_{i=1}^{m} z_i^{\sum_{r=1}^{d} a_{ir}s_r}\right) (\sum_{r=1}^{d} p_r\Pi_{i=1}^{m} z_i^{a_{ir}})^{N-\sum_{r=1}^{d} s_r} \Big\}$$

$$= \frac{N! \prod_{r=1}^{d} p_r^{s_r}}{(N-\sum_{r=1}^{d} s_r)!} \cdot [z_1^{k_1-\sum_{r=1}^{d} a_{1r}s_r} \cdots z_m^{k_m-\sum_{r=1}^{d} a_{dr}s_r}] \Big\{ (\sum_{r=1}^{d} p_r\Pi_{i=1}^{m} z_i^{a_{ir}})^{N-\sum_{r=1}^{d} s_r} \Big\}$$

The multinomial theorem permits the re-writing of the bracketed expression in (†) above:

$$(\ddagger) \quad (\sum_{r=1}^{d} p_r\Pi_{i=1}^{m} z_i^{a_{ir}})^{N-\sum_{r=1}^{d} s_r}$$

$$= \sum_{\substack{(j_1,\ldots,j_d)\in\mathbb{N}^d \\ j_1+\cdots+j_d=N-\sum_{r=1}^{d} s_r}} \binom{N - \sum_{r=1}^{d} s_r}{j_1, \ldots, j_d} \left(\Pi_{r=1}^{d} p_r^{j_r}\right) \left(\Pi_{i=1}^{m} z_i^{\sum_{r=1}^{d} a_{ir}j_r}\right)$$

So that the coefficient of $z_1^{k_1-\sum_{r=1}^{d} a_{1r}s_r} \cdots z_m^{k_m-\sum_{r=1}^{d} a_{dr}s_r}$ in (‡) is:

$$\sum_{\substack{(j_1,\ldots,j_d)\in\mathbb{N}^d \\ j_1+\cdots+j_d=N-\sum_{r=1}^{d} s_r \\ (k_1,\ldots,k_m)^T=\mathbf{A}(j_1+s_1,\ldots,j_d+s_d)^T}} \binom{N - \sum_{r=1}^{d} s_r}{j_1, \ldots, j_d} \left(\Pi_{r=1}^{d} p_r^{j_r}\right)$$

Dividing the above by $[z_1^{k_1} \cdots z_m^{k_m}]G_{\mathbf{Y}}(z_1, \ldots, z_m)$ and multiplying by $\frac{N! \prod_{r=1}^{d} p_r^{s_r}}{(N-\sum_{r=1}^{d} s_r)!}$ will then yield the desired conditional mixed moment.

Acknowledgements

This paper would not have been possible without the support and guidance of Dr. Eduardo Sontag. Thanks are also due to Dr. Doron Zeilberger. Supported in part by grant AFOSR FA9550-11-1-0247.

References

[FS09] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009.

[JKB96] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Discrete multivariate distributions*, John Wiley & Sons, Inc., 1996.

[PW08] R. Pemantle and M. C. Wilson, *Twenty combinatorial examples of asymptotics derived from multivariate generating functions*, SIAM Review **50** (2008), no. 2, 199–272.

[SZ10] E. D. Sontag and D. Zeilberger, *A symbolic computation approach to a problem involving multivariate poisson distributions*, Adv. Appl. Math **44** (2010), no. 4, 359–377.

[Wil94] H. S. Wilf, *Generatingfunctionology*, second ed., Academic Press, Inc., 1994.

[Xek87] E. Xekalaki, *A method of obtaining the probability distribution of m components conditional on l components of a random vector*, Rev. Roumaine Math. Appl. **32** (1987), no. 6, 581–583.