

Eigenanálisis aplicado a diferentes áreas de las ciencias agrícolas y biotecnología: una revisión

Eigenanalysis applied in different areas of agronomical science and biotechnology: a review

Vicente Peña-Caballero¹, Adán Topiltzin Morales-Vargas¹, Carlos Alberto Núñez-Colín^{1*}

¹Programa de Ingeniería en Biotecnología, Universidad de Guanajuato, Correspondencia: Mutualismo 303, esq. Prolongación Río Lerma, col. La Suiza, Apartado Postal 91, 38060, Celaya, Guanajuato, México.

*Autor para correspondencia: lit007a@gmail.com

Fecha de recepción:

6 de septiembre de 2019

Fecha de aceptación:

7 de mayo de 2020

Disponible en línea:

27 de julio de 2020

Este es un artículo en acceso abierto que se distribuye de acuerdo a los términos de la licencia Creative Commons.



Reconocimiento-

NoComercial-CompartirIgual

4.0 Internacional

RESUMEN

El análisis de eigenvalores o eigenanálisis es una técnica matemática, descrita dentro del álgebra lineal, con utilidad en distintas ramas de la ciencia. En este trabajo se hizo la revisión para dos casos específicos, en los cuales esta técnica matemática tiene connotaciones e interpretaciones diferentes. En el primer caso —el análisis de diversidad biológica—, se revisaron los tres análisis factoriales principales: en componentes principales, en coordenadas principales y de correspondencias. En todos se dio una interpretación similar: representar en un plano euclidiano la variabilidad de m unidades experimentales con n variables evaluadas, donde se ocupan tanto los eigenvalores como los eigenvectores. En el segundo caso —la estabilidad de reactores biológicos—, se analizan los eigenvalores de la matriz de estado del biorreactor para definir e interpretar su estabilidad. Esta revisión ofrece una lectura complementaria sobre el uso e interpretación del eigenanálisis para estudios de diversidad biológica y estabilidad de biorreactores.

PALABRAS CLAVE

Eigenvalores, eigenvectores, análisis de diversidad biológica, estabilidad de biorreactores.

ABSTRACT

Analysis of eigenvalues or eigenanalysis is a mathematical technique, which is part of linear algebra, and which can be applied to different branches of the science. This article, reviews two specific cases where this mathematical technique has different connotations and interpretations. In the first case, the biological diversity analysis, the three main factorial analysis were examined: Principal component analysis, principal coordinate analysis, and correspondence analysis. All of these analyses have a similar interpretation: to show the variability of m experimental units with n traits in a Euclidean plane, where both eigenvalues and eigenvectors are used. In the second case, in the stability of biological reactors, the eigenvalues of bioreactor state matrix were analysed to define and to interpret the bioreactor stability. This review offers a complementary reading about the use and interpretation of eigenanalysis for studies on biological diversity and on bioreactors stability.

KEYWORDS

Eigenvalues, eigenvectors, analysis of biological diversity, stability of bioreactors.

INTRODUCCIÓN

El uso de métodos estadísticos multivariados es utilizado en distintas ramas de la ciencia, desde las biológicas (e. g. genética poblacional), las agronómicas (e. g. caracterización de germoplasma), la ingeniería (e. g. estabilidad de reactores), hasta la economía (e. g. econometría) y la sociología (e. g. caracterización de poblaciones). Estos métodos se basan en un solo análisis a partir de un número n de variables de respuesta, medidas en m unidades experimentales, con el objetivo de tratar de comprender sistemas multifactoriales (Johnson 1998); cabe señalar que todo lo anterior es necesario entenderlo a través del álgebra de matrices como elemento básico de análisis (Searle 2006).

Dentro de los métodos multivariados, destacan los análisis basados en eigenanálisis, los cuales reciben su nombre de acuerdo con la matriz inicial que se quiera analizar. Éstos factorizan matrices cuadradas en una matriz con dimensionalidad reducida, lo que permite llegar a conclusiones aun con matrices de datos de gran tamaño (Johnson 1998; Hair et al. 2001).

Con base en lo anterior, en la presente revisión se tiene como objetivo hacer una recapitulación, mediante marcos teóricos y ejemplos, del uso e interpretación que tiene el eigenanálisis en estudios de diversidad biológica —en este caso, llamados análisis factoriales— y en el estudio de estabilidad de biorreactores en biotecnología. Debido a que en artículos recientes en estas áreas se ha observado que existen dudas para la interpretación de la prueba. Además, son dos áreas con interpretaciones contrastantes, lo cual ocasiona que, aunque se tenga el mismo principio matemático, no se obtenga la misma interpretación.

MARCO TEÓRICO

El análisis de eigenvalores o eigenanálisis se basa en la factorización matricial de la matriz $A = [a_{ij}]$, mediante un procedimiento estándar en álgebra lineal, con una matriz A de tamaño $n \times n$, a la cual se le obtiene una ecuación característica al calcular el determinante de la matriz original menos su matriz identidad, también de tamaño $n \times n$, multiplicada por un escalar λ e igualada a cero [1] (Searle 2006).

$$\det(A - \lambda I) = 0 \quad 1$$

Donde $A \in \mathbb{R}^{n \times n}$ es la matriz de coeficientes numéricos o simbólicos del problema en estudio; $I \in \mathbb{R}^{n \times n}$ es la matriz identidad, y λ es un escalar.

Donde las n raíces que resuelven la ecuación [1] (λ) son conocidas como eigenvalores, también llamados valores propios o valores característicos de la matriz A , y donde cada uno de estos eigenvalores (λ_i) está asociado a un eigenvector (u_i), ya sea de filas o de columnas, el cual se calcula de la siguiente manera [2] (Searle 2006).

$$(A - \lambda_i I) u_i = 0 \quad 2$$

Donde $A \in \mathbb{R}^{n \times n}$ es la matriz de coeficientes numéricos o simbólicos del problema en estudio; $I \in \mathbb{R}^{n \times n}$ es la matriz identidad; λ_i es el eigenvalor i de la matriz A , y $u_i \in \mathbb{R}^{n \times 1}$ es el eigenvector del eigenvalor i , ya sea de filas o de columnas.

Los eigenvectores (u_i) forman una matriz $n \times n$, y esos eigenvectores muestran la aportación de cada uno de los elementos de la matriz A al eigenvalor, ya sea de filas o columnas (Searle 2006).

En este trabajo, para ejemplificar este procedimiento se utilizará una matriz de tamaño 2×2 , es decir, $A = [a_{ij}] \in \mathbb{R}^{2 \times 2}$ con $\lambda_i, i = 1, 2$

Suposición 1. Si la matriz A es de rango completo, es decir, sus filas y sus columnas son linealmente independientes, entonces el rango de la matriz A , $\text{rango}(A)$, es igual a n ; en este caso, igual a dos ($\text{rango}(A) = 2$).

Ahora considere el supuesto de que la matriz A [3] es una matriz numérica como sigue:

$$A = \begin{bmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{bmatrix} \triangleq \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \quad 3$$

Y la matriz identidad (I) multiplicada por un escalar λ es

$$\lambda I = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \quad 4$$

Por lo que se aplica la operación $A - \lambda I$

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} a_{11} - \lambda & a_{21} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \quad 5$$

Que, para el ejemplo, quedaría

$$A - \lambda I = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 1 - \lambda & 0.5 \\ 0.5 & 1 - \lambda \end{bmatrix} \quad 5a$$

Acto seguido, se obtiene su determinante

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \det \begin{bmatrix} a_{11} - \lambda & a_{21} \\ a_{12} & a_{22} - \lambda \end{bmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - (a_{12})(a_{21}) \\ &= \lambda^2 - \lambda(a_{11} + a_{22}) + (a_{11})(a_{22}) - (a_{12})(a_{21}) \\ &= \lambda^2 - \lambda \operatorname{tr}z(\mathbf{A}) + \det(\mathbf{A}) \end{aligned} \quad 6$$

Donde se demuestra de [6] que $\det(\mathbf{A}) = a_{11}a_{22} - a_{21}a_{12}$ y $\operatorname{tr}z(\mathbf{A}) = a_{11} + a_{22}$

Para el ejemplo, se obtiene lo siguiente:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^2 - \lambda \operatorname{tr}z(\mathbf{A}) + \det(\mathbf{A}) = \lambda^2 - (1 + 1)\lambda + [(1)(1) - (0.5)(0.5)] = \lambda^2 - 2\lambda + 0.75 \quad 6a$$

Donde, al aplicar la ecuación [1], se obtiene la ecuación característica de la matriz \mathbf{A} ; ahora se resuelve la ecuación [6a], y las raíces que resuelven a λ_i son los eigenvalores, por lo que se puede demostrar fácilmente que las soluciones de λ_1 y λ_2 , obtenidas mediante la siguiente ecuación [7],

$$\lambda_i = \frac{\operatorname{tr}z(\mathbf{A}) \pm \sqrt{\operatorname{tr}z(\mathbf{A})^2 - 4 \cdot \det(\mathbf{A})}}{2}, i = 1, 2 \quad 7$$

al aplicar la ecuación [7], se tienen para este ejemplo las raíces $\lambda_1=1.5$ y $\lambda_2=0.5$.

Ahora, a partir de estos eigenvalores, se obtiene un eigenvector para cada uno de ellos al aplicar la ecuación [2], mediante el siguiente procedimiento.

$$\begin{aligned} (a_{11} - \lambda_i)X_1 + a_{12}X_2 &= 0 \\ a_{21}X_1 + (a_{22} - \lambda_i)X_2 &= 0 \end{aligned} \quad 8$$

Donde u_i es el eigenvector del eigenvalor i (λ_i) y del cual se genera el siguiente sistema de ecuaciones, lo que da la posibilidad de encontrar los valores (X_1 y X_2) de ese eigenvector

$$(\mathbf{A} - \lambda_i \mathbf{I})u_i = \begin{bmatrix} a_{11} - \lambda_i & a_{21} \\ a_{12} & a_{22} - \lambda_i \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad 9$$

Donde X_1 y X_2 son las características de filas o columnas de la matriz \mathbf{A} .

Para el caso de ejemplo, se empezará con $\lambda_1=1.5$, donde, aplicando la ecuación [8], se obtiene lo siguiente

$$(\mathbf{A} - \lambda_1 \mathbf{I})u_1 = \begin{bmatrix} (1 - 1.5) & 0.5 \\ 0.5 & (1 - 1.5) \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad 8a$$

y se alcanza el siguiente sistema de ecuaciones

$$\begin{aligned} -0.5X_1 + 0.5X_2 &= 0 \\ 0.5X_1 - 0.5X_2 &= 0 \end{aligned} \quad 9a$$

el cual, al ser resuelto, se encuentra que

$$u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad 9b$$

Para el caso de $\lambda_2=0.5$, se tiene lo siguiente:

$$(\mathbf{A} - \lambda_2 \mathbf{I})u_2 = \begin{bmatrix} (1 - 0.5) & 0.5 \\ 0.5 & (1 - 0.5) \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad 8b$$

y se obtiene el siguiente sistema de ecuaciones

$$\begin{aligned} 0.5X_1 + 0.5X_2 &= 0 \\ 0.5X_1 + 0.5X_2 &= 0 \end{aligned} \quad 9c$$

Al resolverlo, se encuentra que

$$u_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad 9d$$

Encontrar eigenvalores y eigenvectores tiene diversos usos en las distintas ramas de las ciencias agrícolas —como caracterización de germoplasma, taxonomía numérica, análisis nutrimental, entre otros— y en la biotecnología —estabilidad de reactores, análisis de bifurcación y oscilación de sistemas, control para técnicas de desacoplo, análisis y corrección de imágenes, entre otras—, lo que también se traduce en que tengan diferentes interpretaciones, dependiendo de la matriz \mathbf{A} que se someta a esta factorización y el uso e interpretación específico dado a los eigenvalores y eigenvectores en cada una de estas áreas.

USO E INTERPRETACIÓN DEL EIGENANÁLISIS EN ESTUDIOS DE DIVERSIDAD BIOLÓGICA

Una de las técnicas estadísticas más utilizadas para la caracterización de germoplasma son los análisis factoriales (Núñez-Colín y Escobedo-López 2014), que son diversos métodos ubicados dentro de lo que se conoce como estadística multivariada, y que están

basados en el eigenanálisis (Johnson 1998); éstos son utilizados principalmente para condensar o reducir información contenida en una serie de variables originales en una serie más pequeña de dimensiones compuestas o valores teóricos nuevos (factores), con una mínima pérdida de información, para tener una representación geométrica de las unidades experimentales, las variables originales o ambos, y poder elucidar sus parecidos y diferencias o sus correlaciones (Hair et al. 2001; Núñez-Colín y Valadez-Moctezuma 2010; Núñez-Colín y Escobedo-López 2014).

Los análisis factoriales más utilizados son los descriptivos —llamados así, por no contar con pruebas de hipótesis estadísticas exactas o intervalos de confianza estadísticos— como el Análisis en Componentes Principales (ACP), el Análisis de Coordenadas Principales (ACoP) y el Análisis Factorial de Correspondencias (AFC), que deben utilizarse de acuerdo con el tipo de datos que se evaluarán, principalmente en su naturaleza estadística (Núñez-Colín y Valadez-Moctezuma 2010; Núñez-Colín y Escobedo-López 2014). Aunque estos tres métodos estadísticos tienen su base en el eigenanálisis, la matriz A utilizada en la factorización en eigenvalores y eigenvectores, previamente descrita, es diferente; por lo que se explicará cada uno de ellos por separado, aunque su interpretación tiene el mismo principio (Johnson 1998; Núñez-Colín y Valadez-Moctezuma 2010; Núñez-Colín y Escobedo-López 2014).

ANÁLISIS EN COMPONENTES PRINCIPALES (ACP)

El ACP se basa en factorizar los eigenvalores y los eigenvectores de la matriz de varianzas y covarianzas estandarizadas —matriz de correlaciones momento-producto de Pearson— de n variables de respuesta aplicados a m individuos, siempre y cuando todas las variables tengan o tiendan a tener una distribución normal (Johnson 1998).

La matriz momento-producto de Pearson (ρ) es simétrica, con una diagonal principal con todos sus elementos con valor de 1; por tanto, la traza de la matriz es igual a n ($\text{tr}z(\rho) = n$) y representa la varianza total (Johnson 1998; Searle 2006). Mientras tanto, los otros elementos toman valores entre -1 y 1 ($-1 \leq r_{ij} \leq 1$) (Johnson 1998).

Bajo estos supuestos, la matriz de entrada A queda definida de la siguiente manera

$$\rho = \begin{bmatrix} 1 & r_{21} & \cdots & r_{n1} \\ r_{12} & 1 & \cdots & r_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1n} & r_{2n} & \cdots & 1 \end{bmatrix} \quad 10$$

Donde: $r_{ij} = r_{ji}$ y $-1 \leq r_{ij} \leq 1$.

Al someter a la matriz ρ , de tamaño $n \times n$, a eigenanálisis se obtienen n eigenvalores, todos positivos, los cuales se ordenan de mayor a menor, dándole el nombre de primer componente principal al de mayor eigenvalor; el de segundo, al que le siga en magnitud, y así, sucesivamente, hasta el eigenvalor n (Johnson 1998; Núñez-Colín y Barrientos-Priego 2006; Núñez-Colín y Escobedo-López 2014); la suma de los eigenvalores es la traza de la matriz ($\text{tr}z(\rho) = n$) (Searle 2005) y se le considera la varianza total, por lo que a cada eigenvalor se le obtiene el porcentaje que acumula con respecto a la suma total de los mismos; así, entre más alto sea su eigenvalor, mayor varianza explica. Una vez ordenados los eigenvalores de mayor a menor, y con el porcentaje de varianza aportada por cada componente, se determina la varianza acumulada, para poder explicar la máxima variación en pocas dimensiones (Núñez-Colín y Barrientos-Priego 2006; Núñez-Colín y Escobedo-López 2014). El análisis es más preciso cuando se acumula más de 50% de la varianza total al tercer componente, donde se puede tener una representación bi o tridimensional (Núñez-Colín y Barrientos-Priego 2006).

Los eigenvectores asociados (u_i) a cada eigenvalor en el ACP se correlacionan con las variables originales para conocer cuáles de ellas están relacionadas con cada componente principal y determinar las variables principales, que son aquellas de mayor variación. Estos eigenvectores (u_i) se utilizan para hacer una proyección de las m unidades experimentales o individuos, utilizando las siguientes consideraciones:

La matriz básica de datos M es una matriz con m individuos y n variables de respuesta, donde cada individuo es un vector (I_i) con n elementos y_{jk} ; a esta matriz es a la que se le obtiene ρ , de tamaño $n \times n$, con n eigenvalores, cada uno de ellos con su eigenvector asociado (u_i) con elementos x_{ik} ; después, se multiplican

los vectores para obtener el valor de proyección del *i*ésimo individuo en el *i*ésimo componente principal

$$I_j \cdot u'_i = [y_{j1} \quad y_{j2} \quad \dots \quad y_{jn}] \begin{bmatrix} x'_{i1} \\ x'_{i2} \\ \vdots \\ x'_{in} \end{bmatrix} = \sum_{k=1}^n y_{jk} \cdot x'_{ik} \quad 11$$

Con estos valores, los individuos son proyectados en planos de dimensionalidad reducida, para buscar el parecido con base en la distancia que generen entre ambos puntos.

Para ejemplificar el uso de esta técnica, se utilizarán datos de una caracterización bioquímica de 12 variedades de sorgo blanco (*Sorghum bicolor* L. Moench), mediante 6 variables bioquímicas (Chiquito-Almanza et al. 2016). La matriz de correlación (ρ) se describe en el Cuadro 1.

A esta matriz se le sacan sus eigenvalores utilizando las ecuaciones [1] y [7], donde la suma de estos eigenvalores es la traza de ρ ($\text{tr}(\rho)=n$). La correlación [10] se ve como una matriz de varianzas y covarianzas estandarizada, por lo que la diagonal principal de la matriz representa las varianzas estandarizadas de cada una de las variables. Cabe recordar que un dato estandarizado se obtiene de la siguiente manera:

$$Z_i = \frac{(x_i - \bar{x})}{S_x} \quad 12$$

Donde Z_i es el valor estandarizado de x_i ; x_i es el *i*ésimo valor de la variable x ; \bar{x} es el valor promedio de la variable x , y S_x es la desviación estándar de la variable x .

Lo que hace que cada variable tenga una distribución normal estándar, que presenta media igual a cero y varianza igual a uno ($\bar{x} = 0$ y $S_x^2 = 1$), por lo que la diagonal principal son las varianzas estandarizadas —con todos los valores igual a 1— y los demás datos las covarianzas estandarizadas; de esta manera, el análisis parte de que las variables tienen el mismo peso en la varianza final de ρ .

Al ordenar los componentes principales (CP), que están ordenados por su máxima variación, se ve la importancia de cada uno de ellos, con base en la proporción de varianza que explican. En este caso, el primer CP explica 40.9786% de la varianza total; el segundo explica 30.2027% y el tercero 21.3312%. Entre los tres, acumulan 92.5125%, por lo que es la explicación que puede verse en un plano tridimensional, es decir, lo que ojo humano puede interpretar de fácil manera (Cuadro 2).

Cuadro 1. Matriz de correlaciones (ρ) de doce genotipos de sorgo dulce mediante seis características bioquímicas (Chiquito-Almanza et al. 2016).

Variable	Proteína total	Rendimiento de kafirina	Proporción de kafirina	Rendimiento de lisina	Proporción de lisina	Digestibilidad
Proteína total	1.0000	0.6076	-0.5585	-0.0815	-0.4778	-0.1672
Rendimiento de kafirina	0.6076	1.0000	0.3124	-0.0804	-0.3338	0.1448
Proporción de kafirina	-0.5585	0.3124	1.0000	0.0043	0.2154	0.3811
Rendimiento de lisina	-0.0815	-0.0804	0.0043	1.0000	0.9131	-0.4881
Proporción de lisina	-0.4778	-0.3338	0.2154	0.9131	1.0000	-0.3719
Digestibilidad	-0.1672	0.1448	0.3811	-0.4881	-0.3719	1.0000

Cuadro 2. Eigenvalores de la matriz de correlaciones (ρ) de doce genotipos de sorgo dulce mediante seis características bioquímicas (Chiquito-Almanza et al. 2016).

CP	Eigenvalor (λ_i)	Proporción de la varianza	Varianza acumulada
1	2.4587	40.9786	40.9786
2	1.8122	30.2027	71.1813
3	1.2799	21.3312	92.5125
4	0.4444	7.4067	99.9192
5	0.0039	0.0643	99.9836
6	0.0010	0.0165	100.0000
Suma	6.0000	100.0000	

Donde la proporción de la varianza es calculada como $\%S_i^2 = \left(\frac{\lambda_i}{\sum_{i=1}^n \lambda_i}\right) \cdot 100$.

Acto seguido, se calculan los eigenvectores de cada CP, teniendo lo siguiente para el ejemplo (Cuadro 3).

Cuadro 3. Eigenvectores de los tres primeros Componentes Principales (CP) de doce genotipos de sorgo dulce mediante seis características bioquímicas (Chiquito-Almanza et al. 2016).

Variable	CP1	CP2	CP3
Proteína total	0.5760	0.7463	0.2902
Rendimiento de kafirina	0.5255	0.0960	0.8424
Proporción de kafirina	-0.1290	-0.8070	0.5280
Rendimiento de lisina	-0.8240	0.3195	0.3999
Proporción de lisina	-0.9646	-0.0120	0.2175
Digestibilidad	0.4741	-0.7017	-0.0050

En negritas están los valores con mayores correlaciones entre el CP y las variables originales

Los eigenvectores deben entenderse como la aportación de cada variable a dicho componente, y se interpretan como una correlación; es decir, valores iguales o cercanos a cero significan que no hay relación significativa entre esta variable y el componente; valores cercanos o iguales a 1, significan una correlación positiva, y valores cercanos o iguales a -1 significan una correlación negativa. No importa qué tipo de correlación tenga, por lo que se seleccionan valores cercanos o iguales al valor absoluto de uno (|1|).

En el ejemplo, las variables rendimiento y proporción de lisina son las más importantes en el CP1 y presentan una correlación negativa. En el CP2, fueron importantes las variables proteína total (correlación positiva), proporción de kafirinas y digestibilidad (ambas con una correlación negativa). Finalmente, en el CP sólo fue importante la proporción de kafirinas, con una correlación positiva.

Posteriormente, se genera una proyección de los genotipos en los primeros tres CP, al multiplicar el eigenvector por los datos de cada genotipo (aplicando la

ecuación [11]); se debe tener cuidado de utilizar los datos estandarizados de cada genotipo, para evitar sesgos por las unidades en las que está cada variable. Para el ejemplo, se tiene lo siguiente (Figura 1).

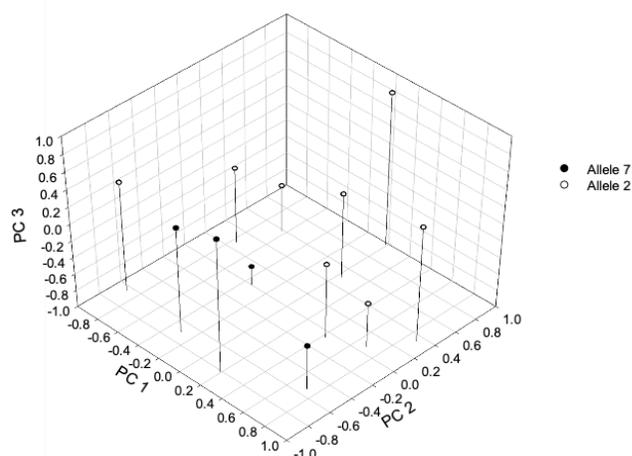


Figura 1. Proyección de los doce genotipos de sorgo dulce en los primeros tres Componentes Principales (elaborada con base en Chiquito-Almanza et al. 2016).

La interpretación de este gráfico (Figura 1) es que entre más cercanos estén los puntos, más parecidos son los genotipos, y entre más alejados, más diferentes. Si están en el lado positivo del CP1 (en nuestro ejemplo con correlaciones negativas), tendrán bajo valor en la proporción y rendimiento de lisina; ocurre lo contrario con los que se encuentren ubicados en la parte negativa del CP1. Si la correlación fuera positiva —como en el caso del CP3 de nuestro ejemplo—, significa que si están en la parte positiva del CP3, el valor de la proporción de kafirinas fue alto, y si están ubicadas en la parte negativa del CP3, entonces presenta bajos valores de proporción de kafirinas.

ANÁLISIS EN COORDENADAS PRINCIPALES (ACooP)

El Análisis en Coordenadas Principales se obtiene al generar los eigenvalores de una matriz de índices de similitud, la cual se centraliza con la distancia de Gower (1966); los eigenvectores de dicho análisis sirven como coordenadas de las unidades experimentales en planos euclidianos, y las variables sólo son para obtener los índices de similitud, a diferencia del Análisis en Componentes Principales. Por tanto, este análisis sólo sirve para tener una representación de las unidades experimentales y comparar su comportamiento con otros análisis clasificatorios como el análisis clúster.

Los índices de similitud —matriz de entrada de este análisis— siguen el siguiente principio descrito

por Núñez-Colín y Escobedo-López (2011). Suponga dos unidades experimentales a las que se les ha sacado datos binarios; las combinaciones de comparación posibles se muestran en el Cuadro 4.

Cuadro 4. Tabla de posibilidades y codificación base de los índices de similitud utilizados en caracterización de germoplasma.

	Individuo <i>j</i>	
	1	0
Individuo <i>i</i>	1	0
	<i>a</i>	<i>b</i>
	0	<i>d</i>
	<i>c</i>	

Donde *a* es que los dos individuos tengan presente la misma característica; *b* es que el primero la presente y el segundo no; *c* que el segundo presente la característica y el primero no, y *d* que ambos carezcan de esta característica.

A partir de esto, se calcula *m*, que es la suma entre *a* y *d* ($m = a + d$), es decir, la suma de las concordancias, tanto de presencia como de ausencia; *u*, que es la suma de *b* y *c* ($u = b + c$), es decir, la suma de las discordancias, y *n*, que es el número total de características evaluadas, es decir, la suma de las concordancias y las discordancias ($n = m + u$). Estas codificaciones se utilizan en la actualidad, a nivel mundial, en los artículos relacionados con la teoría y el desarrollo de estos índices; a partir de esta codificación, tenemos las siguientes fórmulas (Cuadro 5).

Cuadro 5. Índices de similitud más utilizados.

Nombre del Coeficiente	Fórmula	Autor
Simple Matching*	$\frac{m}{n}$	Sokal y Sneath (1963)
Jaccard	$\frac{a}{n - d}$	Jaccard (1908)
Dice	$\frac{2a}{2a + b + c}$	Dice, (1945); Nei y Li (1979)
Phi	$\frac{(ad - bc)}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$	Sokal y Sneath (1963)
Kulczynski 1	$\frac{a}{u}$	Kulczynski (1927)
Kulczynski 2	$\frac{1}{2} \left[\left(\frac{a}{a + b} \right) + \left(\frac{a}{a + c} \right) \right]$	Kulczynski (1927)
Hamann*	$\frac{m - u}{m}$	Hamann (1961)
Roger y Tanimoto*	$\frac{n}{(n + u)}$	Rogers y Tanimoto (1960)
Russel y Rao	$\frac{a}{n}$	Russel y Rao (1940)
Ochiai	$\frac{a}{\sqrt{(a + b)(a + c)}}$	Ochiai (1957)
Yule	$\frac{(ad - bc)}{(ad + bc)}$	Yule (1911)

* Estas fórmulas aceptan datos cualitativos nominales.

Cada uno de estos índices fueron desarrollados para fines específicos; por ejemplo, Simple matching (Sokal y Sneath 1963) fue desarrollado para analizar características evaluadas a partir de claves taxonómicas dicotómicas, en donde un ejemplar podía o no tener ciertas características y en raros casos se presentaban más de dos posibilidades de respuesta; ésta es la razón por la cual este índice sí permite datos cualitativos nominales. Sin embargo, si lo que se va a analizar son matrices binarias de presencia/ausencia de fragmentos de huellas genómicas, obtenidas por electroforesis, el índice más adecuado es la fórmula desarrollada por Dice (1945). Aunque creado originalmente con fines de comparación de nichos ecológicos, Nei y Li (1979) demostraron —mediante desarrollos matemáticos— que era el índice más congruente para la clasificación de este tipo de características, debido principalmente a que una concordancia por doble ausencia de un fragmento puede no deberse a la misma razón, es decir, no es una concordancia (Núñez-Colín y Escobedo-López 2011).

La matriz obtenida es simétrica en tamaño $m \times m$, donde m es el número de unidades experimentales en estudio, y los valores dentro de la matriz (a_{ij}) están en el rango de valor entre 0 y 1, por lo que es necesario centralizarla —o doble centralizarla para ser preciso—, para lo cual se ejecuta la siguiente fórmula:

$$\delta_{ij} = a_{ij} - \bar{a}_i - \bar{a}_j + \bar{a} \quad 13$$

Donde δ_{ij} son los elementos de la i ésima columna y la j ésima fila de la matriz doble centralizados; a_{ij} son los elementos de la i ésima columna y la j ésima fila de la matriz; \bar{a}_i es el valor promedio de la i ésima columna de la matriz; \bar{a}_j es el valor promedio de la j ésima fila de la matriz, y \bar{a} es el promedio general de todos los valores de la matriz.

Estos valores pueden ser positivos o negativos, lo que —según Gower (2005)— hace una representación euclidiana aproximada al someterse a eigenanálisis, donde los eigenvectores funcionan como las coordenadas de cada unidad experimental en un plano bidimensional o tridimensional; se considera a la primera coordenada la de mayor eigenvalor (eje X), y a la segunda la siguiente en eigenvalor (eje Y), y así sucesivamente. A diferencia del ACP, en el ACooP la proporción de cada uno de los eigenvalores en relación con la suma total de los mismos no es tan importante como lo es la comparación de resultados con análisis clasificatorios, como el Análisis Clúster, donde se tiene una representación en un plano euclidiano de las

unidades experimentales y se observa el parecido entre ellas mediante la distancia que presentan.

ANÁLISIS DE CORRESPONDENCIAS (AFC)

El Análisis Factorial de Correspondencias, o simplemente Análisis de Correspondencias, se hace a partir de matrices rectangulares $n \times m$; al no ser matrices cuadradas, es necesario hacer una transformación en una matriz de Burts, es decir, una multiplicación de la matriz por su matriz transpuesta para poder obtener los eigenvalores.

$$\mathbf{A}_{Burts} = \mathbf{A} \cdot \mathbf{A}^T \quad 14$$

Donde $\mathbf{A} \in \mathbb{R}^{n \times m}$ y $\mathbf{A}_{Burts} \in \mathbb{R}^{n \times n}$ o $\mathbf{A}_{Burts} \in \mathbb{R}^{m \times m}$ dependiendo de si m o n otorgan el menor rango a la matriz.

Pero, para los eigenvectores, éstos se basan en la obtención de una distancia χ^2 —en lugar de la distancia euclidiana—, en nubes de puntos tanto de fila como de columna (Montenegro y Pardo 1996).

La distancia euclidiana [15] y la distancia χ^2 [16] se calculan mediante las siguientes expresiones:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ki} - x_{kj})^2} \quad 15$$

$$d_{ij} = \sqrt{\sum_{k=1}^n \frac{1}{x_k} \left(\frac{x_{ki}}{x_i} - \frac{x_{kj}}{x_j} \right)^2} \quad 16$$

Donde d_{ij} es la distancia entre los individuos i y j ; x_{ki} es el dato de la variable k ésima en el individuo i ésimo; x_k es el promedio de la variable k ésima en todas las unidades experimentales, y x_i es el promedio de todas las variables en el individuo i ésimo.

Al AFC se le denomina simple o múltiple, dependiendo de si se basa en la tabla de contingencia de dos variables con números efectivos (simple) o de si se fundamenta en datos de muchas variables de respuesta y muchas unidades experimentales (Montenegro y Pardo 1996). Al aplicar la ecuación [16], se debe cuidar que todos los datos estén en las mismas escalas, para evitar sesgos, pues por la fórmula de la distancia χ^2 [16] no es posible estandarizarla, porque los promedios por variable (X_k) serían cero (Núñez-Colín et al. 2004; Núñez-Colín y Valadez-Moctezuma 2010).

Este análisis tiene exactamente la misma interpretación que se le da al Análisis en Componentes Principales, pero, en lugar de proyecciones, se tienen

datos de eigenvectores para filas y para columnas (variables y unidades experimentales), lo que permite tener una menor cantidad de cálculos.

No obstante, la interpretación de este análisis es totalmente análoga al ACP, sólo que, en lugar de eigenvectores para las variables y proyecciones con variables estandarizadas, en el AFC se ocupan los eigenvectores para variables y para unidades experimentales en las nubes punto de filas y de columnas, según correspondan.

EL EIGENANÁLISIS EN ESTUDIOS DE ESTABILIDAD DE REACTORES BIOLÓGICOS

En un contexto general, la biotecnología a través del uso de microorganismos —por ejemplo, células eucariotas y procariontes— aprovecha sus capacidades metabólicas para generar bienes y servicios. El cultivo de los microorganismos se desarrolla en los llamados biorreactores, reactores biológicos o fermentadores. El éxito del cultivo de algún microorganismo está relacionado con la estabilidad de operación del reactor, es decir, que opere en condiciones óptimas de transferencia de materia y energía. En la práctica, para abstraer el bioproceso de bioconversión en el biorreactor se desarrolla un modelo matemático; éste puede ser de base fenomenológica, a través de la aplicación de un balance clásico de materia y/o energía sobre el reactor (Peña-Caballero et al. 2012; 2016). En este sentido, disponer de un modelo que describa el bioproceso es una herramienta importante para evaluar las propiedades del proceso en el biorreactor. Una propiedad importante para evaluar la correcta operación del biorreactor es su propiedad de estabilidad; esta propiedad se puede evaluar matemáticamente mediante el análisis de los eigenvalores de la matriz de estado del proceso A .

Por lo general, para este tipo de estudios, la matriz A —de la ecuación [1]— se construye a partir del conjunto de ecuaciones diferenciales ordinarias no lineales para n variables de interés, que modelan el proceso en estudio mediante la aplicación de balances clásicos de materia y energía; es decir, matemáticamente resulta que:

$$\begin{aligned} \frac{d}{dt} \mathbf{x} &= \mathbf{f}(\mathbf{x}, \mathbf{u}), \mathbf{x}(0) = \mathbf{x}_0, \\ \mathbf{y} &= \mathbf{g}(\mathbf{x}, \mathbf{u}) \end{aligned} \quad 17$$

Donde $\mathbf{x} \in \mathbb{R}^{n \times 1}$ es el vector de variables o variables de estado, que son el conjunto de variables que describen el sistema de estudio;

$\mathbf{x}(0) \in \mathbb{R}^{n \times 1}$ es el vector de estados iniciales en $t = 0$; $\mathbf{y} \in \mathbb{R}^{r \times 1}$ es el vector de variables medibles u observables, y $\mathbf{u} \in \mathbb{R}^{m \times m}$ es el vector de variables de entrada o de control.

LINEALIZACIÓN LOCAL ALREDEDOR DE UN PUNTO DE EQUILIBRIO

Definición 1 (Equilibrio). Un par $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \in \mathbb{R}^{n \times 1} \times \mathbb{R}^{m \times 1}$ se llama un punto de equilibrio de [17] si $\mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}}) = 0$. En este caso, $\mathbf{u}(t) = \bar{\mathbf{u}}, \mathbf{x}(t) = \bar{\mathbf{x}}, \mathbf{y}(t) = \bar{\mathbf{y}} := \mathbf{g}(\bar{\mathbf{x}}, \bar{\mathbf{u}}), \forall t \geq 0$ es una solución de la ecuación [17].

Al linealizar la función $\mathbf{f}(\cdot)$ en la ecuación [17] mediante una serie de expansión de Taylor, truncada para el término lineal, se obtiene la versión linealizada del modelo en la ecuación [17].

Definición 2 (Linealización local alrededor de un punto de equilibrio). El sistema lineal de la ecuación [17] es:

$$\begin{aligned} \frac{d}{dt} \mathbf{x} &= \mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u} \\ \mathbf{y} &= \mathbf{C} \mathbf{x} + \mathbf{D} \mathbf{u} \end{aligned} \quad 18$$

Definido por las siguientes matrices jacobianas evaluadas en el punto de equilibrio $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$

$$\mathbf{A} := \frac{\partial \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}})}{\partial \mathbf{x}}, \mathbf{B} := \frac{\partial \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{u}})}{\partial \mathbf{u}}, \mathbf{C} := \frac{\partial \mathbf{g}(\bar{\mathbf{x}}, \bar{\mathbf{u}})}{\partial \mathbf{x}}, \mathbf{D} := \frac{\partial \mathbf{g}(\bar{\mathbf{x}}, \bar{\mathbf{u}})}{\partial \mathbf{u}}$$

es llamado la linealización local de [17] alrededor de un punto de equilibrio (\mathbf{x}, \mathbf{u}) .

Donde $\mathbf{A} = [a_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{n \times n}$; $\mathbf{B} = [b_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{n \times m}$; $\mathbf{C} = [c_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{r \times n}$; $\mathbf{D} = [d_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{r \times m}$; $\mathbf{x} = [x_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{n \times 1}$; $\mathbf{y} = [y_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{r \times 1}$, y $\mathbf{u} = [u_{ij}]: [0, \infty) \rightarrow \mathbb{R}^{m \times 1}$, son la matriz de estados, matriz de control, matriz de salida, matriz de transición de control, vector de estados, vector de salida y vector de control, respectivamente.

Ahora, en los bioprocesos de conversión de substratos a productos por un microorganismo confinado en un biorreactor (ecuación [19]), el análisis de la propiedad de estabilidad del biorreactor es una parte importante para propósitos de control y calidad del proceso. En un sentido general, la estabilidad del biorreactor se puede analizar a través de sus valores propios o eigenvalores en el sentido de Lyapunov (Khalil 2002). En este apartado, se evalúa la estabilidad de un biorreactor sulfato reductor [19] mediante la técnica de estabilidad puntual del criterio de estabilidad de Lyapunov.

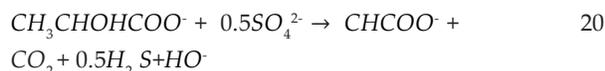
$$\mathbf{X} + \mathbf{S} \overset{\mathbf{X}}{\rightarrow} \mathbf{X} + \mathbf{P} \quad 19$$

Donde \mathbf{X}, \mathbf{S} y \mathbf{P} son biomasa, substrato y producto, respectivamente.

PROCESO SULFATO REDUCTOR

Las bacterias sulfato reductoras son microorganismos anaerobios que pueden crecer en presencia de diferentes fuentes de carbono y energía, como alcoholes, hidrógeno, lactato, entre otras fuentes (Feio et al. 2004; Kamarisima et al. 2019; Nielsen et al. 2019; Spring et al. 2019). *Desulfovibrio vulgarius* y *D. desulfuricans* son las bacterias sulfato reductoras más ampliamente estudiadas por su capacidad para remover y oxidar metales pesados (Ayangbenro et al. 2018; Li et al. 2018; Ravikumar et al. 2018; Serrano y Leiva 2017) al aplicarlas en procesos de biorremediación. En años recientes, una nueva bacteria fue aislada e identificada como sulfato-reductora, *D. alaskensis* (Feio et al. 2004), la cual ha despertado interés de la comunidad científica e industrial para su aplicación al diseño de nuevas tecnologías de biorremediación de aguas contaminadas con efluentes industriales. Más recientemente, Neria-González et al. (2006) aislaron e identificaron la bacteria *D. alaskensis* 6SR, en un ducto para petróleo del sureste de México. La bacteria presentó alta capacidad para crecer en medio Postgate C modificado (Postgate 1984), con presencia de altas concentraciones de sulfato y metales pesados. En este sentido, la *D. alaskensis* 6SR se ha empleado en la remoción de metales pesados, por ejemplo, cadmio (López-Pérez et al. 2013) y cromo hexavalente (Peña-Caballero et al. 2016), dos de los contaminantes colocados en la lista de prioridad por la Organización Mundial de la Salud (OMS), por su efecto negativo sobre los seres vivos y el medio ambiente.

Por lo anterior, en esta sección se considera un bioproceso de remoción de sulfato en medio líquido, utilizando como sistema biológico la bacteria *D. alaskensis* 6SR (Neria-González et al. 2006). *Desulfovibrio alaskensis* 6SR es una bacteria estrictamente anaerobia, la cual utiliza el sulfato SO_4^{2-} , como aceptor final de electrones, al oxidar al lactato $CH_3CHOHCOO^-$, de acuerdo con la siguiente reacción estequiométrica:



En los cultivos de *D. alaskensis* 6SR, la presencia de los iones de sulfuro puede ser un factor limitante del bioproceso, volviéndolo inestable en el biorreactor, debido a que el sulfuro de hidrógeno inhibe el crecimiento de la bacteria. Por ejemplo, en Peña-Caballero et al. (2016), se propone el modelado del proceso de *D. alaskensis* 6SR mediante una cinética no estructurada tipo Levenspiel, que considera en su estructura la concentración del sulfuro de hidrógeno P :

$$\frac{\mu(S, P)}{\mu_{max}} = \frac{S}{K_M + S} \left[1 - \frac{P}{P^*} \right]^n \quad 21$$

Donde P es la concentración de sulfuro de hidrogeno en el volumen de reacción en el biorreactor; P^* es la concentración de inhibición de sulfuro de hidrógeno en un dominio de 500-560 mg/L para *D. alaskensis* 6SR a $D = 0.01 - 0.025$ 1/h con lactato como fuente de carbono, y n es una constante que se relaciona con la inhibición del producto (Han y Levenspiel 1988).

La elevada producción de sulfuro de hidrógeno H_2S en el biorreactor, en ambas fases: líquido y gas, es decir, en equilibrio termodinámico, ha motivado el estudio de la dinámica de H_2S en las dos fases, para su aplicación en la remoción de cromo hexavalente Cr(VI), a través del diseño de nuevos sistemas de producción de H_2S , y su uso para remover Cr(VI) en procesos controlados (Peña-Caballero et al. 2016).

DINÁMICA DEL PROCESO SULFATO REDUCTOR

La dinámica del proceso sulfato reductor se puede modelar considerando la estequiometría en la ecuación [20] y como se muestra en la cinética de reacción en la ecuación [21]. En este trabajo, el proceso se modela considerando únicamente la concentración de biomasa y sulfato (ver reacción en la ecuación [20]).

Para propósitos de análisis de estabilidad en el sentido de Lyapunov, se considera un modelo reducido para *D. alaskensis* 6SR (Peña-Caballero et al. 2012). Brevemente, el modelo estudiado considera dos variables de estado: biomasa X [g/L] y substrato S [g/L], con una cinética de tipo Monod (Monod 1949), considerando como substrato limitante al sulfato (ver ecuación [21]). El proceso opera en modo continuo, es decir, el volumen de reacción V [L] se mantiene constante en el tiempo y en estado estacionario (Figura 2).

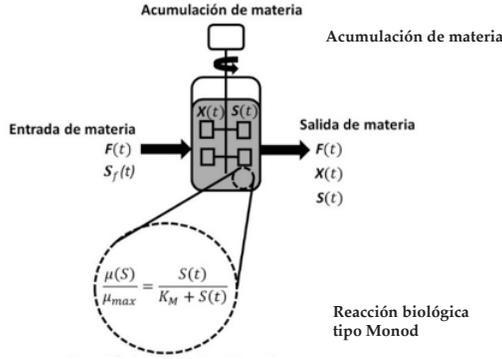


Figura 2. Biorreactor operando en continuo con una cinética de crecimiento celular tipo Monod, considerando un sustrato limitante (quimioestado).

La cinética de crecimiento celular de *D. alaskensis* 6SR se define como $\mu = f(S)$ [1/h] y es de tipo Monod (Monod 1949). La representación del proceso en función de la ecuación [17] y la ecuación [18] son las ecuaciones [22] y [23], respectivamente.

$$\begin{aligned} \frac{dX}{dt} &= (\mu - D)X \\ \frac{dS}{dt} &= D(S_f - S) - \frac{\mu}{Y}X \end{aligned} \quad (22)$$

$$\mathbf{y} = h\left(\begin{bmatrix} X \\ S \end{bmatrix}\right) = c\left[\begin{bmatrix} X \\ S \end{bmatrix}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ S \end{bmatrix}$$

$$\frac{d}{dt} \begin{bmatrix} X \\ S \end{bmatrix} = \begin{bmatrix} (\bar{\mu} - \bar{D}) & \dot{\mu}\bar{X} \\ -\frac{\bar{\mu}}{Y} & -\bar{D} - \frac{\dot{\mu}}{Y}\bar{X} \end{bmatrix} \begin{bmatrix} X \\ S \end{bmatrix} + \begin{bmatrix} -\bar{X} & 0 \\ (\bar{S}_f - \bar{S}) & \bar{D} \end{bmatrix} \begin{bmatrix} \bar{D} \\ \bar{S}_f \end{bmatrix} = \mathbf{0}$$

$$\mathbf{y} = \begin{bmatrix} \bar{X} \\ \bar{S} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{X} \\ \bar{S} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{D} \\ \bar{S}_f \end{bmatrix} \quad (23)$$

Donde \bar{X} y \bar{S} son las concentraciones de biomasa y sustrato en estado estacionario, es decir, $\frac{d}{dt} \begin{bmatrix} \bar{X} \\ \bar{S} \end{bmatrix} = \mathbf{0}$ si \bar{D} es la tasa de dilución en estado estacionario; \bar{S}_f es la concentración de sustrato en estado estacionario en la corriente de alimentación; Y es el rendimiento de biomasa a sustrato consumido, y $\frac{d(\cdot)}{dt} = (\dot{\cdot})$ la variación de la propiedad (\cdot) en función de la variación del tiempo.

De la ecuación [23] se deduce que la matriz $A(\cdot)$ es:

$$\mathbf{A} = \begin{bmatrix} (\bar{\mu} - \bar{D}) & \dot{\mu}\bar{X} \\ -\frac{\bar{\mu}}{Y} & -\bar{D} - \frac{\dot{\mu}}{Y}\bar{X} \end{bmatrix} \quad (24)$$

Como se mencionó, la cinética es de tipo Monod (Monod 1949):

$$\mu(S) = \mu_{max} \frac{S}{K_M + S} \quad (25)$$

En la ecuación [25], la derivada del modelo cinético de Monod (Monod 1949) en estado estacionario μ es:

$$\dot{\mu}(\bar{S}) = \frac{\mu_{max} K_M}{(K_M + \bar{S})^2} \quad (26)$$

En la ecuación [24], se puede demostrar de \mathbf{A} que:

$$\det(\mathbf{A}) = (\bar{\mu} - \bar{D}) \left(-\bar{D} - \frac{\dot{\mu}}{Y}\bar{X} \right) + \frac{\bar{\mu}}{Y} (\dot{\mu}\bar{X}) \quad (27)$$

$$\text{trz}(\mathbf{A}) = (\bar{\mu} - \bar{D}) - \bar{D} - \frac{\dot{\mu}}{Y}\bar{X} \quad (28)$$

Pero en estado estacionario se cumple la siguiente igualdad:

$$(\bar{\mu} - \bar{D}) = 0 \quad (29)$$

por lo que las ecuaciones [27] y [28] toman la siguiente forma:

$$\det(\mathbf{A}) = \frac{\dot{\mu}\bar{X}\bar{\mu}}{Y} \quad (30)$$

$$\text{trz}(\mathbf{A}) = -\bar{\mu} - \frac{\dot{\mu}\bar{X}}{Y} \quad (31)$$

Y , aplicando las definiciones de las ecuaciones [30] y [31] en la ecuación [7], se puede demostrar que:

$$\lambda_1 = \bar{\mu} \text{ [1/h]} \text{ y } \lambda_2 = \frac{\dot{\mu}}{Y}\bar{X} \text{ [1/h]} \quad (32)$$

Ahora, para ejemplificar el uso de los valores propios en la ecuación [32], para analizar la estabilidad del biorreactor con la bacteria sulfato reductora, se considera la operación del reactor para diferentes condiciones de operación en la ecuación [22]; es decir, se alcanzan diferentes puntos de equilibrio (\bar{X}, \bar{S}) . Específicamente, cada estado de equilibrio se obtiene para diferentes tasas de dilución $D = F/V$ (ver Figura 2). En el cuadro 6, se presentan cuatro diferentes valores para D . Para las diferentes tasas de dilución, les corresponden diferentes estados de equilibrio de las variables del proceso \bar{X} y \bar{S} , es decir, los estados estacionarios (ver cuadro 6). En la figura 3 se puede observar que, para todos los casos, la biomasa y el sustrato alcanzan un nuevo estado de equilibrio, por lo que sus concentraciones son diferentes. Los eigenvalores para cada estado de equilibrio se presentan en el cuadro 6. Los valores de λ_1 y λ_2 permiten decidir dónde no operar el reactor para evitar esta condición de operación. El comportamiento dinámico de los estados para cada valor de D se muestran en la figura 4

para las diferentes tasas de dilución. Los experimentos numéricos se desarrollaron considerando los siguientes datos: $S_f = 5 \frac{g}{L}$, $Y = \frac{g_{biomasa}}{g_{substrato}}$, $\mu_{max} = 0.035 \text{ 1/h}$, $K_M = 0.9 \text{ g/L}$, $D = [0-0.025] \text{ 1/h}$. Con las siguientes condiciones iniciales: $X(t_0) = 0.12 \text{ g/L}$ y $S(t_0) = 5 \text{ g/L}$ (Peña-Caballero et al. 2012).

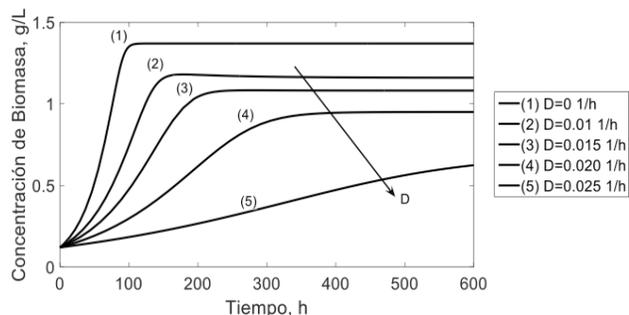


Figura 3. Dinámica de la concentración de biomasa para diferentes tasas de dilución.

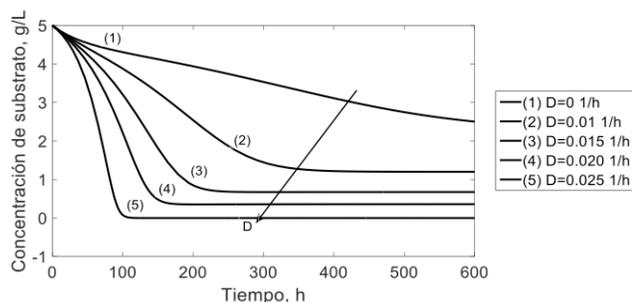


Figura 4. Dinámica de la concentración de sustrato para diferentes tasas de dilución.

Cuadro 6. Estados estacionarios y valores propios del modelo en la ecuación [3] para diferentes tasas de dilución.

$D \text{ [1/h]}$	X	S	λ_1	λ_2
0.010	1.1603	0.3599	-0.0250	-0.1071
0.015	1.0813	0.6750	-0.0250	-0.0649
0.020	0.9500	1.2001	-0.0250	-0.0321
0.025	0.6242	2.5031	-0.0060	-0.0250

Para evaluar el tipo de estabilidad del reactor en cada punto de equilibrio en el cuadro 6, se recurre a la estabilidad de Lyapunov.

ESTABILIDAD DE LYAPUNOV

Considere el sistema en [17]

Definición 3 (Estabilidad de Lyapunov). El sistema en la ecuación [17] es:

1. *(Marginalmente) estable* en el sentido de Lyapunov o internamente estable si, para cada condición inicial $x(t_0) = x_0 \in \mathbb{R}^{n \times 1}$, la respuesta del estado homogéneo es uniformemente acotada.

$$x(t) = \Phi(t, t_0)x_0, \quad \forall t \geq 0 \quad 33$$

2. *Asintóticamente estable* (en el sentido de Lyapunov) si, adicionalmente para cada condición inicial $x(t_0) = x_0 \in \mathbb{R}^{n \times 1}$, tenemos que $x(t) \rightarrow 0$ como $t \rightarrow \infty$.
3. *Exponencialmente estable* si, además, existe la constante $c, \lambda > 0$ tal que, para cada condición inicial $x(t_0) = x_0 \in \mathbb{R}^{n \times 1}$, tenemos

$$\|x(t)\| \leq c \exp(\lambda(t - t_0)) \|x(t_0)\|, \quad \forall t \geq 0 \quad 34$$

4. *Inestable*, si no es marginalmente estable en el sentido de Lyapunov. Las matrices $B(\cdot)$, $C(\cdot)$ y $D(\cdot)$ no desempeñan un papel en las definiciones anteriores; solamente importa $A(\cdot)$ porque define completamente la matriz de transición de estados Φ .

Condiciones de valor propio para estabilidad de Lyapunov. El sistema en la ecuación [18] es:

1. *Marginalmente estable* si y sólo si al menos uno de los eigenvalores de A tiene parte real cero y el resto parte real negativa.
2. *Asintóticamente estable* si y solamente si todos los eigenvalores de A tienen estrictamente parte real negativa, pero con al menos un valor que tiende a cero.
3. *Exponencialmente estable* si y solamente si todos los eigenvalores de A tienen estrictamente parte real negativa.
4. *Inestable* si y solamente si al menos uno de los eigenvalores de A tiene parte real positiva (Hespanha 2009). Entonces, considerando los resultados de los valores de los eigenvalores del cuadro 7, se puede concluir que los diferentes estados de equilibrio que se alcanzan en el biorreactor para los diferentes valores de D son asintóticamente estables. Además, los resultados de estabilidad se reafirman, debido a que el determinante de la matriz A no es cero y para todos los casos de operación del biorreactor la matriz A es de rango completo e igual a n , es decir, las filas y las columnas son linealmente independientes.

Cuadro 7. Estabilidad en el sentido de Lyapunov para diferentes estados estacionarios para la dinámica del biorreactor en la ecuación [30] para diferentes tasas de dilución.

D [1/h]	$\det(A)$	rango (A)	λ_1	λ_2	Estabilidad en el sentido de Lyapunov
0.010	0.0027	2	-0.0250	-0.1071	asintóticamente estable
0.015	0.0016	2	-0.0250	-0.0649	asintóticamente estable
0.020	8.0349e-04	2	-0.0250	-0.0321	asintóticamente estable
0.025	1.5119e-04	2	-0.0060	-0.0250	asintóticamente estable

Finalmente, de la ecuación [7] se puede desprender el cuadro para las diferentes posibilidades de los eigenvalores de λ_1 y λ_2 . Se puede precisar que para los diferentes estados de equilibrio del biorreactor cumplen que $\sqrt{\text{tr}z(A)^2 - 4 \cdot \det(A)} > 0$, es decir, $\lambda_i \in \mathbb{R}$; $\lambda_1 \neq \lambda_2$ (Cuadro 8).

Cuadro 8. Solución de [8]: tipos de eigenvalores y la forma de las soluciones.

$\sqrt{\text{tr}z(A)^2 - 4 \cdot \det(A)}$	λ_i	Solución
> 0	λ_1 y λ_2	$\lambda_i \in \mathbb{R}$; $\lambda_1 \neq \lambda_2$
= 0	λ_1 y λ_2	$\lambda_i \in \mathbb{R}_-$; $\lambda_1 = \lambda_2$
< 0	λ_1 y λ_2	$\lambda_i \in \mathbb{C}$; $\lambda_1 = a + bi$; $\lambda_2 = a - bi$

COMENTARIOS FINALES

Como se puede constatar en esta revisión, el eigenanálisis es útil para diferentes estudios dentro de las ciencias agrícolas y la biotecnología. No obstante, cada variante de la aplicación de esta técnica matemática lleva a tener diferentes connotaciones e interpretaciones, por lo que debe comprenderse cada uno de los procesos que se están llevando a cabo para poder hacer una interpretación correcta de los resultados y no cometer errores de interpretación por creer que, por tener el mismo principio, deben ser lo mismo.

AGRADECIMIENTOS

Esta investigación fue financiada por la Universidad de Guanajuato, mediante el proyecto CIIC 006/2019.

LITERATURA CITADA

- Ayangbenro AS, Olanrewaju OS, Babalola OO. 2018. Sulfate-reducing bacteria as an effective tool for sustainable acid mine bioremediation. *Frontiers in Microbiology* 9: Article 1986. <https://doi.org/10.3389/fmicb.2018.01986>
- Chiquito-Almanza E, Ochoa-Zarzosa A, López-Meza JE, Pecina-Quintero V, Núñez-Colín CA, Anaya-López JL. 2016. A new allele of γ -kafirin gene coding for a protein with high lysine content in Mexican white sorghum germplasm. *Journal of the Science of Food and Agriculture* 96(10): 3342-3350. <https://doi.org/10.1002/jsfa.7513>
- Dice LR. 1945. Measure of the amount of ecologic associations between species. *Ecology* 26(3): 277-302. DOI: 10.2307/1932409
- Feio MJ, Zinkevich V, Beech IB, Brossa EL, Eaton P, Schmitt J, Guezennec J. 2004. *Desulfovibrio alaskensis* sp. nov., a sulphate reducing bacterium from a soured oil reservoir. *International Journal of Systematic and Evolutionary Microbiology* 54: 1747-1752. <https://doi.org/10.1099/ijs.0.63118-0>
- Gower JC. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53(3-4): 325-338. <https://doi.org/10.1093/biomet/53.3-4.325>
- Gower JC. 2005. Principal Coordinates Analysis. In: Armitage P, Colton T, editors. *Encyclopedia of Biostatistics*. Chichester, John Wiley & Sons.
- Hair JF Jr, Anderson RE, Tatham RL, Black WC. 2001. *Análisis multivariante*. Prentice Hall Iberia. Madrid, España.
- Hamann U. 1961. Merkmalbestand und Verwandtschaftsbeziehungen der Farinosae. Ein Beitrag zum System der Monokotyledonen. *Willdenowia* 2(5): 639-768.
- Han K, Levenspiel O. 1988. Extended Monod kinetics for substrate, product and cell inhibition. *Biotechnology and Bioengineering* 32(4): 430-447. <https://doi.org/10.1002/bit.260320404>
- Hespanha JP. 2009. *Linear Systems Theory*. Princeton University Press. Princeton, USA.
- Jaccard P. 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles* 44(163): 223-270. <https://doi.org/10.5169/seals-268384>
- Johnson DE. 1998. *Métodos multivariados aplicados al análisis de datos*. International Thomson Editores. Ciudad de México, México.
- Kamarisima, Miyanaga K, Tanji Y. 2019. The utilization of aromatic hydrocarbon by nitrate- and sulfate-reducing bacteria in single and multiple nitrate injection for souring control. *Biochemical Engineering Journal* 143: 75-80. <https://doi.org/10.1016/j.bej.2018.12.006>
- Khalil HK. 2002. *Nonlinear Systems*. Prentice Hall. New Jersey, USA.
- Kulczynski S. 1927. Die Pflanzenassoziationen der Pienienen. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres, Classe des Sciences Mathematiques et Naturelles B (Suppl. 2)*: 57-203.
- Li YC, Yang XY, Geng B. 2018. Preparation of immobilized sulfate-reducing bacteria-microalgae beads for effective bioremediation of copper-containing wastewater. *Water Air and Soil Pollution* 229(3): article 54. <https://doi.org/10.1007/s11270-018-3709-1>
- López-Pérez PA, Neria-González MI, Aguilar-López R. 2013. Cadmium concentration stabilization in a class of continuous sulfate reducing bioreactor via sulfide concentration control. *Chemical Papers* 67(3): 326-335. <https://doi.org/10.2478/s11696-012-0274-8>
- Monod J. 1949. The growth of bacterial cultures. *Annual Reviews in Microbiology* 3(1): 371-394. <https://doi.org/10.1146/annurev.mi.03.100149.002103>
- Montenegro A, Pardo CE. 1996. *Introducción al análisis de datos textuales*. Universidad Nacional de Colombia. Santa Fe, Colombia.
- Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science (USA)* 76(10): 5269-5273.
- Neria-González I, Wang ET, Ramírez F, Romero JM, Hernández-Rodríguez C. 2006. Characterization of bacterial community associated to biofilms of corroded oil pipelines from the Southeast of Mexico. *Anaerobe* 12(3): 122-133. <https://doi.org/10.1016/j.anaerobe.2006.02.001>
- Nielsen G, Coudert L, Janin A, Blais JF, Mercier G. 2019. Influence of organic carbon sources on metal removal from mine impacted water using sulfate-reducing bacteria bioreactors in cold climates. *Mine Water and the Environment* 38(1): 104-118. <https://doi.org/10.1007/s10230-018-00580-3>

- Núñez-Colín CA, Barrientos-Priego AF. 2006. Estimación de la variabilidad interna de muestras poblacionales, mediante el análisis en componentes principales. *Interciencia* 31(11): 802-806.
- Núñez-Colín CA, Valadez-Moctezuma E. 2010. Análisis Estadístico de Huellas Genómicas, un Uso Práctico de los Paquetes Computacionales más Populares. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. Celaya, Guanajuato.
- Núñez-Colín CA, Escobedo-López D. 2011. Uso correcto del análisis clúster en la caracterización de germoplasma vegetal. *Agronomía Mesoamericana* 22(2): 415-427.
- Núñez-Colín CA, Escobedo-López D. 2014. Caracterización de germoplasma vegetal: la piedra angular en el estudio de los recursos fitogenéticos. *Acta Agrícola y Pecuaria* 1(1): 1-6.
- Núñez-Colín CA, Rodríguez-Pérez JE, Nieto-Ángel R, Barrientos-Priego AF. 2004. Construcción de dendrogramas de taxonomía numérica mediante el coeficiente de distancia χ^2 : una revisión. *Revista Chapingo Serie Horticultura* 10(2): 229-237. <https://doi.org/10.5154/r.rchsh.2003.07.046>
- Ochiai A. 1957. Zoogeographic studies on the soleoid fishes found in Japan and its neighbouring regions. *Bulletin of the Japanese Society of Scientific Fisheries* 22(9): 526-530. <https://doi.org/10.2331/suisan.22.526>
- Peña-Caballero V, López-Pérez PA, Neria-González MI, Aguilar-López R. 2012. A class of nonlinear adaptive controller for a continuous anaerobic bioreactor. *Journal of Scientific & Industrial Research* 71(7): 480-486.
- Peña-Caballero V, Aguilar-López R, López-Pérez PA, Neria-González MI. 2016. Reduction of Cr(VI) utilizing biogenic sulfide: an experimental and mathematical modeling approach. *Desalination and Water Treatment* 57(28): 13056-13065. <https://doi.org/10.1080/19443994.2015.1055811>
- Postgate JR. 1984. *The Sulphate-Reducing Bacteria*. (2nd edition). Cambridge University Press. Cambridge, UK.
- Ravikumar KVG, Argulwar S, Sudakaran SV, Pulimi M, Chandrasekaran N, Mukherjee A. 2018. Nano-Bio sequential removal of hexavalent chromium using polymer-nZVI composite film and sulfate reducing bacteria under anaerobic condition. *Environmental Technology & Innovation* 9: 122-133. <https://doi.org/10.1016/j.eti.2017.11.006>
- Rogers DG, Tanimoto TT. 1960. A computer program for classifying plants. *Science* 132(3434): 1115-1118. <https://doi.org/10.1126/science.132.3434.1115>
- Russel PF, Rao TR. 1940. On habitat and association of species of Anopheline larvae in south-eastern Madras. *Journal of the Malaria Institute of India* 3(1): 153-178.
- Searle SR. 2005. Eigenvalue. In: Armitage P, Colton T, editors. *Encyclopedia of Biostatistics*, 6. Chichester, John Wiley & Sons.
- Searle SR. 2006. *Matrix Algebra Useful for Statistics*. Wiley-Interscience. New Jersey, USA.
- Serrano J, Leiva E. 2017. Removal of arsenic using acid/metal-tolerant sulfate reducing bacteria: a new approach for bioremediation of high-arsenic acid mine waters. *Water* 9(12): 994. <https://doi.org/10.3390/W9120994>
- Sokal RR, Sneath PHA. 1963. *Principles of Numerical Taxonomy*. H. Freeman & Company. San Francisco, USA.
- Spring S, Sorokin DY, Verbarq S, Rohde M, Woyke T, Kyrpidis NC. 2019. Sulfate-reducing bacteria that produce exopolymers thrive in the calcifying zone of a hypersaline cyanobacterial mat. *Frontiers in Microbiology* 10: a862. <https://doi.org/10.3389/Fmicb.2019.00862>
- Yule GU. 1911. *An Introduction of the Theory of Statistics*. Charles Griffin & Company. London, UK.