

RESUMEN: Herramientas Matemáticas para el Análisis de Actividad Física *

Oscar Eduardo Martínez Castiblanco Universidad Sergio Arboleda

Resumen: *Este resumen no incluye gráficas ni tablas.*

El auge de la actividad física post pandemia ha estado ligado a las aplicaciones y dispositivos que permiten hacer seguimiento de las mismas. Esto sumado a la popularidad que ganaron algunos deportes que favorecen el distanciamiento social (ciclismo, atletismo) da origen a una gran cantidad de datos para ser analizados. En este trabajo se analiza la topografía del Alto de Guasquita, ubicado a 40 km de la ciudad de Bogotá, con base en datos recolectados a través de un GPS. Aplicando conceptos trigonométricos, de cálculo y estadística en R, se obtiene una comprensión de la estructura del puerto, y sugerencias para optimizar el desempeño deportivo de ciclistas que transitan esta zona.

Keywords: Ciencia de datos, Estadística, Cálculo, Trigonometría Esférica, Deporte, Ciclismo, Matemática Aplicada, R, XML, knitr, plotly, pracma, dplyr, forecast, cluster

Introducción

La bicicleta hace parte de la vida y las rutinas de muchos colombianos; desde el niño que recibe una como regalo de navidad hasta el gran campeón que conquista las rutas europeas, pasando por quien se gana la vida trabajando en ella. Este trabajo tiene sus orígenes en el ciclismo aficionado, entendido como la práctica esporádica de este deporte con fines recreativos o de salud. Esta práctica, combinada con la proliferación de dispositivos GPS y redes sociales, han dado origen a una gran cantidad de información obtenida a través de colaboración abierta distribuida (crowdsourcing). Estos datos han sido utilizados en trabajos académicos que permiten entender tanto el desempeño individual (Wolf et al., 2018) (Repici, 2019) (Wamsley, 2014) como las tendencias grupales de los ciclistas urbanos y recreativos (Pogodzinska, Kiec and D'Agostino, 2020) (Fischer, Nelson and Winters, 2022) (Sun and Mobasher, 2017). En este trabajo, se toman datos generados en competencia, recolectados con un ciclocomputador, y descargados de la red social Strava para analizar el perfil del Alto de Guasquita, dar una interpretación adecuada a su pendiente, y ponderar su dificultad (desde lo deportivo) con respecto a otros puertos de la zona con pendiente similar.

Las matemáticas del ciclismo

La bicicleta es un vehículo de propulsión humana. La fuerza de las piernas se aplica a dos bielas a través de pedales y hacen girar una rueda dentada conocida como plato. El plato transfiere dicha fuerza a una rueda dentada más pequeña llamada piñón, que se encuentra sujeta a la rueda trasera, y de esta manera fuerza aplicada se convierte en energía cinética, la rueda gira

* Autor de contacto: oscar.martinez@usa.edu.co.

y se produce el movimiento. La eficiencia de la bicicleta se debe a las relaciones existentes entre los tamaños del plato y del piñón.

Sea D el diámetro del plato, d el diámetro del piñón, F la frecuencia con la que el plato gira y f la frecuencia con la que el piñón girara. La velocidad angular con la que gira el plato está dada por $2\pi F$ y al aplicar la ecuación fundamental de velocidades para transmisiones por correa y despejar f se obtiene $f = \frac{FD}{d}$. Teniendo en cuenta que $D > d$, se concluye que $f > F$, es decir, la frecuencia con la que gira el piñón es mayor a la del plato, en otras palabras, una vuelta del plato genera más de 1 vuelta del piñón y por ende, más de 1 vuelta de la rueda trasera. Sin pérdida de generalidad, la relación de transmisión $\frac{D}{d}$ se puede interpretar en términos de los dientes de cada engranaje, ya que tienen el mismo tamaño tanto en el plato como en el piñón. Si N es la cantidad de dientes del plato, y n la del piñón, la razón de marchas está dada entonces por $\frac{N}{n}$, y se interpreta como las vueltas que da el piñón por cada vuelta del plato. Estas vueltas se multiplican por el diámetro de la rueda, obteniendo así la distancia recorrida por cada vuelta completa de los pedales. Cabe anotar que la fuerza necesaria para poner en marcha el sistema es directamente proporcional a las razones previamente mencionadas. En otras palabras: a mayor razón se genera mayor avance, y se requiere más fuerza para poner en marcha el sistema.

La inclinación de las vías se calcula con la fórmula de la pendiente de una recta, y se interpreta de manera porcentual. En términos prácticos, los dispositivos GPS permiten obtener longitud, latitud y altitud, lo que permite calcular la pendiente. En ese sentido, si entre coordenadas planas (longitud, latitud) se avanzan 100m y se gana 1m de elevación, la vía tiene una inclinación del 1%. El problema radica ahora en calcular la distancia recorrida dadas dos coordenadas planas. Para tal fin, se utiliza la fórmula de Haversine ([Brummelen, 2013](#)). Esta expresión permite calcular la distancia entre dos puntos $P = (x_{lat}, x_{lon})$ y $q = (y_{lat}, y_{lon})$ sobre el globo terraqueo, a partir de sus coordenadas. Esta fórmula se puede calcular en R a través de la función `haversine` dentro de la librería `pracma`. Una vez calculadas las distancias horizontales, y teniendo los datos de altitud, es posible visualizar la ruta, así como calcular las inclinaciones de la misma punto a punto, y obtener una mayor comprensión de la vía.

Contexto: Alleycat race

Los datos a ser modelados en este trabajo corresponden a la carrera ciclista “XChallenge”. Dicha carrera se realizó el 26 de febrero de 2022 con partida en el Alto de Patios (La Calera, Cundinamarca). Esta carrera ha sido la única realizada en esta zona con la modalidad de “Alleycat Race”. Al momento de la inscripción, los participantes solo conocían la distancia y desnivel positivo acumulado a recorrer, y que el tiempo de los ganadores se determinaría en cuatro “X” (puertos de montaña cronometrados) cuya ubicación solo será conocida durante la carrera ([Gómez and Aris-tizabal, 2022](#)). El día de la competencia, los participantes se dirigían al punto de partida donde recibían la indicación para llegar a la primera X (alto cronometrado), su longitud y porcentaje de inclinación. Al finalizar el primer puerto cronometrado, los participantes reciben la indicación del siguiente para dirigirse allí y continuar la carrera. Teniendo en cuenta el tipo de competencia, la interpretación que el competidor hace de la distancia y porcentaje de inclinación del puerto (en muchos casos desconocido) es fundamental para diseñar una estrategia de carrera óptima que permita recorrer cada una de las “X” en el menor tiempo posible.

La segunda “X” era el Alto de Guasquita, con 2.7 km y una inclinación de 5.2%. Si bien es cierto

que el porcentaje de inclinación es considerable, este es mucho menor que el primero. Con base en la longitud del puerto y la experiencia en otros ascensos con inclinación similar era plausible pensar en una relación de transmisión entre 2.5 y 3, que con una cadencia de 60 rpm permitiera realizar el ascenso tomando entre 7 y 9 minutos. Sin embargo, los primeros metros del ascenso evidenciaron una dificultad mucho más alta de la que indicaba el porcentaje de inclinación, lo que motivó a realizar un análisis detallado del puerto, para tener una mayor comprensión de su dificultad, no explicada totalmente por el porcentaje de inclinación.

Procesamiento y análisis de datos

Una vez finalizada la carrera y por ende, el registro de los datos, se descargan dos archivos a través de la aplicación Strava, con extensiones gpx y fit. Dada la naturaleza de los archivos, la librería XML permite extraer la información requerida de este tipo de archivos a un DataFrame. Con este conjunto de datos es posible realizar varias visualizaciones de la ruta. Entre ellas, el recorrido de la vía. Además, es posible calcular las distancias recorridas entre cada punto, y así realizar el análisis de la pendiente. También es posible visualizar la vía en 3D utilizando el paquete plotly.

Utilizando la ecuación ?? con la función haversine sobre las columnas de latitud y longitud es posible obtener la distancia entre cada par de coordenadas, y con vectores de diferencias diff se pueden obtener las variaciones de altitud y tiempo entre puntos. Con esta información es posible calcular la velocidad e inclinación en cada punto. Aunque todas las medidas están en metros y segundos, la velocidad se expresa en kilómetros por hora en aras de facilitar su interpretación.

Con estos datos es posible obtener un primer acercamiento a la pendiente de la montaña, realizando el cociente entre la suma de las elevaciones y las distancias punto a punto. Este primer dato que se obtiene es el que aparece referenciado en la aplicación Strava.

Resultados

Con los datos procesados es posible visualizar y analizar la pendiente, en aras de comprender mejor la inclinación de la montaña. El perfil permite evidenciar algunos tramos de menor inclinación, casi planos, lo cual quiere decir que otros tramos deben tener pendiente que excede el promedio. Para comprender mejor este concepto, es posible interpretar la pendiente como una señal. Para clasificar las pendientes, se asignan cuatro categorías: $[0, 2)$, $[2, 6)$, $[6, 10)$, $[10, 40)$. Estos grupos se plantean agrupando pendientes con dificultad similar. Estos porcentajes distribuidos a lo largo de la vía dan cuenta de las tensiones del recorrido, con curvas de alta pendiente seguidas de segmentos lineales cortos de aparente descanso.

```
## Warning: package 'forecast' was built under R version 4.3.2
```

```
## Registered S3 method overwritten by 'quantmod':  
##   method             from  
##   as.zoo.data.frame zoo
```

Finalmente, para comprender cómo se relacionan las pendientes entre ellas, y generar una agrupación que permita comparar el puerto con otros de pendiente similar, se utiliza el algoritmo

k-means sobre la inclinación, distancia y velocidad en cada punto. Con base en el método del codo se generan 4 grupos sobre los datos. Los grupos etiquetados 3 (azul) y 2 (verde) corresponden a segmentos de baja inclinación, y se diferencian principalmente por la velocidad alcanzada en los mismos. Esta velocidad depende de la longitud del tramo y de la inclinación anterior, es decir, el grupo 1 representa tramos más largos que el tramo 2, lo que permite descansar y aumentar la velocidad, mientras que en el grupo 2 los tramos son más cortos, y aunque permiten descansar, su cercanía con un tramo de pendiente mayor no permite que la velocidad aumente como en el grupo anterior. Por otra parte, los grupos 1 (amarillo) y 4 (rojo) corresponden a los tramos de mayor dificultad debido a su pendiente, lo que disminuye el avance en los mismos. Esta información permite identificar dos tramos al final del puerto donde, de acuerdo a la categorización obtenida, se podría aumentar la potencia para obtener mejoras significativas en el tiempo de ascenso.

Los grupos etiquetados 3 (azul) y 2 (verde) corresponden a segmentos de baja inclinación, y se diferencian principalmente por la velocidad alcanzada en los mismos. Esta velocidad depende de la longitud del tramo y de la inclinación anterior, es decir, el grupo 1 representa tramos más largos que el tramo 2, lo que permite descansar y aumentar la velocidad, mientras que en el grupo 2 los tramos son más cortos, y aunque permiten descansar, su cercanía con un tramo de pendiente mayor no permite que la velocidad aumente como en el grupo anterior.

Finalmente, para ponderar la dificultad del Alto de Guasquita, se analiza otro puerto de la zona con una inclinación promedio similar: el Alto del Neusa. Ubicado a 80 km del Alto de Guasquita, cuenta con 6,9 km de longitud y una inclinación promedio del 5.2%. Estos valores se obtienen aplicando el mismo procesamiento de datos presentado previamente sobre archivos .gpx y .fit. Posterior a ello, se clasifica el Alto del Neusa con los centroides generados con el Alto de Guasquita para contrastar los dos recorridos. A diferencia de Guasquita, Neusa tiene un alto porcentaje de recorrido con inclinaciones cercanas a la inclinación promedio. En otras palabras, el porcentaje promedio de inclinación en el Alto del Neusa describe mejor la dificultad del mismo, permitiendo planificar estrategias de potencia y cadencia constantes para el ascenso.

En conclusión, aunque ambos puertos de montaña tengan un porcentaje de inclinación similar, la topografía de los mismos hace que las estrategias de ascenso generadas a partir de uno de ellos no sean del todo aplicables al otro. Además, el análisis desarrollado permite entender la estructura del ascenso con el fin de identificar tramos relevantes al diseñar estrategias de competencia cuyo propósito sea minimizar el tiempo en dicho recorrido.

References

- Brummelen, Glen Van. 2013. *Heavenly Mathematics: The Forgotten Art of Spherical Trigonometry*. Princeton University Press.
- Fischer, Jaimy, Trisalyn Nelson and Meghan Winters. 2022. "Riding through the pandemic: Using Strava data to monitor the impacts of COVID-19 on spatial patterns of bicycling." *Transportation Research Interdisciplinary Perspectives* 15:100667.
URL: <https://www.sciencedirect.com/science/article/pii/S2590198222001270>
- Gómez, Sergio and Juan Aristizabal. 2022. *Guía del Atleta X Challenge*. TNC Games.
- Pogodzinska, Sylwia, Mariusz Kiec and Carmelo D'Agostino. 2020. "Bicycle Traffic Volume Estimation Based on GPS Data." *Transportation Research Procedia* 45:874–881. Transport Infrastructure and systems in a changing world. Towards a more sustainable, reliable and smarter mobility. TIS Roma 2019 Conference Proceedings.
URL: <https://www.sciencedirect.com/science/article/pii/S2352146520301319>
- Repici, Michael. 2019. Because My Garmin Told Me To: A New Materialist Study of Agency and Wearable Technology Phd thesis University of South Florida South Florida, FL: . Available at <https://www.proquest.com/openview/ce940cfa2eed18380a7476f7085543ee/1?pq-origsite=gscholar&cbl=18750&diss=y>.
- Sun, Yeran and Amin Mobasher. 2017. "Utilizing Crowdsourced Data for Studies of Cycling and Air Pollution Exposure: A Case Study Using Strava Data." *International Journal of Environmental Research and Public Health* 14(3).
URL: <https://www.mdpi.com/1660-4601/14/3/274>
- Wamsley, Kyle. 2014. Optimal power-based cycling pacing strategies for Strava segments Phd thesis Kutztown University of Pennsylvania Pennsylvania, PA: . Available at <https://www.proquest.com/openview/17c398b22550ef77615a6b5dfcdc89fa/1?pq-origsite=gscholar&cbl=18750>.
- Wolf, Stefan, Alexander Artiga Gonzalez, Raphael Bertschinger and Dietmar Saupe. 2018. "Modeling in Road Cycling for Optimal Pacing Strategies: Theory vs. Practice." 7:18–19.
URL: <https://www.jsc-journal.com/index.php/JSC/article/view/397>