CdA and Crr from data

Friday, October 5, 2018

Mark:

Apologies for this crude way of communicating -- it was so much fun to talk with you via video call this morning. Anyway, here's what I was thinking.

We can decompose our familiar power equation into four parts: power demand from rolling resistance, climbing (changes in potential energy), acceleration (changes in kinetic energy) and aero drag.

$$W = W_{RR} + W_{PE} + W_{KE} + W_{AERO}$$

$$= Crr \cdot m \cdot g \cdot v + s \cdot m \cdot g \cdot v + mav + \frac{1}{2} \rho \, CdA \, v_a^2 \, v$$

If there's no wind, airspeed = ground speed so va = v and the last term just has a v^3 in it.

Anyway, for any segment of the ride from time i to time j in seconds we can integrate the watts up to joules.

$$J_{ik} = Crr \cdot m \cdot g \cdot \int_i^k v \, dt + m \cdot g \int_i^k s v \, dt + m \int_i^k a v \, dt + \frac{1}{2} \rho \, CdA \int_i^k v^3 \, dt$$

$$= Crr \cdot m \cdot g \; \text{distance}\Big|_i^k + m \cdot g \cdot \text{height}\Big|_i^k + m \cdot \frac{1}{2} v^2 \Big|_i^k + CdA \, \frac{1}{2} \rho \int_i^k v^3 \, dt$$

$$J_{ik} - m \cdot g \cdot \text{height}\Big|_i^k - \frac{m}{2} v^2 \Big|_i^k = Crr \cdot m \cdot g \cdot \text{distance}\Big|_i^k + CdA \, \frac{1}{2} \rho \int_i^k v^3 \, dt$$

AND FINALLY,

$$\underbrace{\frac{J_{ik} - m \cdot g \cdot \text{height}\Big|_i^k - \frac{m}{2} v^2 \Big|_i^k}{m \cdot g \cdot \text{distance}\Big|_i^k}}_{Y} = Crr + CdA \underbrace{\left[\frac{\frac{1}{2} \rho \int_i^k v^3 \, dt}{m \cdot g \cdot \text{distance}\Big|_i^k} \right]}_{X}$$

$$Y = Crr + CdA \quad X$$

So, if we regress Y on X, the intercept will be Crr and the slope will be CdA. This is true for ALL segments from time i to k.

If you're on a 250m velodrome, and i and k are chosen to be the amount of time it takes to go 250m, we know the distance will be 250m and the height difference is zero. So the only issue is keeping track of how many seconds it takes to go 250m: if you're fast that might be 16 seconds; if you're slow that might be twice as long. The length of each interval varies but the elevation change is fixed (at zero). With your data, you observe v each second, so you can get v^2 and v^3 for the kinetic energy and aero terms.

If you're not on a velodrome, you need a way to estimate the change in height over each interval. Suppose you had an (accurate) altimeter. If it were accurate, you could get the change in height over each i to k interval. The advantage of the velodrome is that we know the net elevation change each lap is zero; on the open road we need one more bit of information from another sensor: an altimeter. If so, we don't need to use intervals of variable length -- if the altimeter were accurate, we could fix the intervals to, say, 30 or 60 seconds and let the elevation gain vary. We've traded variable intervals and fixed elevation for fixed intervals and variable elevation -- but at the cost of needing another accurate input. After an initial period of 30 or 60 seconds, each second later we could update the calculation. This gives "real-time" updates of CdA and Crr each second, though they would be the average CdA and average Crr over the preceding 30 or 60 seconds. *If* you also knew that you were doing laps you could either do fixed distance (and go back to having zero net elevation) or do fixed time intervals and variable elevation. That's kind of why I like doing laps.

This, however, is noisy. We can do slightly better.

We can do better if we can put bounds on Crr. Crr in the regression above is estimated by the intercept, so it's "outside" of the range of observed speeds. We can trade off some of the variance in the CdA estimate at the cost of potential bias in the Crr estimate by fixing the Crr or at least restricting it to a narrow range, which could be reasonable if the road surface is fairly uniform.

The other thing that improves the regression is if we get a wide range in speeds and powers. You may recall that if the data are collinear, you can't identify the coefficients. You may also recall that if the data aren't collinear but are *nearly* so, then the standard errors on the coefficients will be large so although the overall fit may be good the individual coefficients are not known with precision. So the best way to "widen" the range of the data is to be sure to vary speed and pwer -- that's why I usually throw in a few coasting bits, even when going on the flat or on slight uphills. That ensures that the estimated coefficients will be more stable and better estimated.

Finally, of course, when we're outdoors we need to account for cases when airspeed isn't equal to ground speed. That's when we have to include the possibility of yet another damn sensor.