Erik van Zwet
Mathematisch Instituut
Universiteit Leiden
Niels Bohrweg 1
2333 CA Leiden
evanzwet@math.leidenuniv.nl

Vakantiecursus 2003

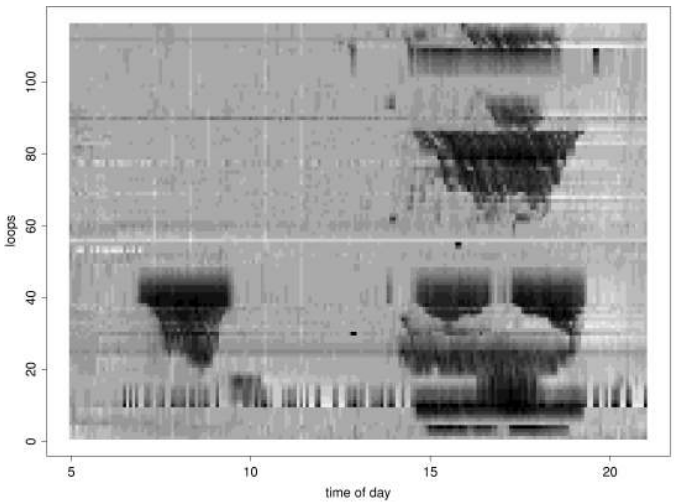
Reistijden voorspellen op snelwegen

Het fileprobleem staat al jaren in de belangstelling van de media en politiek (zie ook het Nieuw Archief voor Wiskunde van juni 2003). De Nederlandse snelwegen kunnen het steeds maar groeiende aantal auto's niet verwerken en de individuele weggebruiker kan zijn bestemming vaak niet meer op tijd bereiken. Het probleem wordt door deskundigen dikwijls als onoplosbaar aangemerkt. In dit artikel betoogt Erik van Zwet dat het voorspellen van de reistijd de onzekerheid van het op tijd arriveren voor een groot deel kan wegnemen. Erik van Zwet is als statisticus verbonden aan de Universiteit Leiden. Hij verzorgde in de zomer van 2003 colleges over het voorspellen van reistijden op snelwegen in het kader van de door het CWI georganiseerde vakantiecursus voor wiskundeleraren.

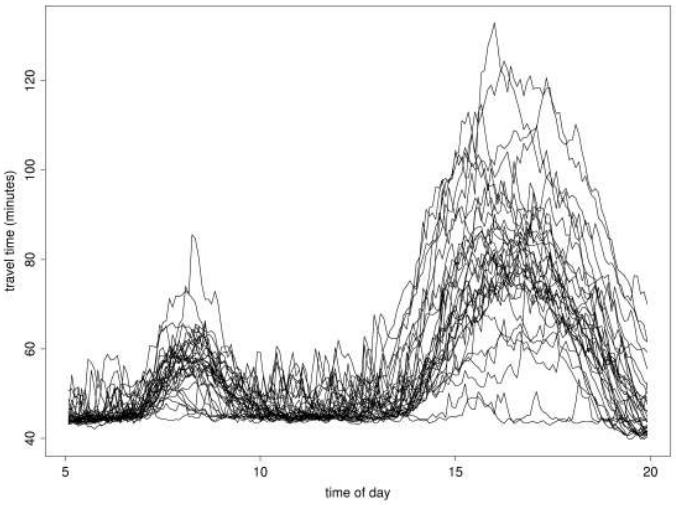
Op bijna alle snelwegen in Nederland liggen zogeheten lusdetectoren. Deze detectoren tellen het aantal passerende voertuigen, en meten hun snelheid. Hetzelfde soort metingen wordt ook gedaan in Californië, waar ik de afgelopen drie jaar onderzoek heb gedaan.

In figuur 1 zien we snelheidsmetingen uit Californië op 16 juni 2000 van 116 lussen langs ongeveer 80 km van Interstate 10 (East) door Los Angeles. Op de horizontale as staat de tijd van de dag, van 's ochtends 5 uur tot 's avonds 9 uur, en op de verticale as staat het lusnummer van 1 tot en met 116. De grijswaarde geeft de snelheid weer bij een bepaalde lus op een bepaalde tijd van de dag. Hoe donkerder de grijswaarde, hoe lager de snelheid. In de donkere, driehoekige vormen herkennen we files die ontstaan en weer oplossen. Kennelijk is het op dit traject 's middags drukker dan 's ochtends.

Met de lusgegevens kunnen we ook een ruwe berekening maken van de reistijd tussen een tweetal punten A en B. In figuur 2 zien we een grafiek van de reistijden op ons traject langs I10 East op weekdagen tussen 16 juni en 8 september 2002. De ochtend- en middagspits zijn duidelijk te herkennen. Op 3 en 4 juli, feestdagen



Figuur 1 Snelheidsveld v op 16 juni 2000 langs Interstate 10 East in Los Angeles. De donkere, driehoekige vormen zijn files die ontstaan en weer oplossen. De horizontale strepen zijn slecht werkende detectoren.



Figuur 2 Reistijden op 34 weekdagen op een traject van 80 km door Los Angeles. Opvallend zijn de grote verschillen in reistijd, vooral in de middagspits.

in Amerika, zijn de reistijden veel korter dan normaal.

Op basis van reistijden uit het verleden kunnen we voorspellen hoe lang het op een gegeven moment in de toekomst zal duren om van A naar B te komen.

Als u of ik ergens naar toe gaan, proberen we natuurlijk te voorspellen wanneer we aan zullen komen. We gebruiken daarbij onze ervaring uit het verleden en luisteren naar de verkeersinformatie op de radio. Onze voorspeller doet in zekere zin hetzelfde. De *ervaring* van de voorspeller is een databestand met de lusdetectormetingen langs de route op elke minuut van de dag in de afgelopen maanden. De voorspeller maakt een combinatie van deze historische gegevens en de meest recente lusmetingen van, zeg, een paar minuten geleden. Om de optimale combinatie te bepalen maken we gebruik van een statistische methode die (lineaire) regressie heet. Deze term komt van de uitdrukking *regressie naar het gemiddelde*. In de praktijk betekent het dat als de verkeerssituatie uitzonderlijk slecht is, we verwachten dat het beter zal worden en vice versa. Hoe snel de verbetering of verslechtering in de richting van het gemiddelde zal plaatsvinden hangt af van verschillende factoren, zoals de locatie en de tijd van de dag.

Bepaalde routes zijn nu eenmaal erg druk, en daardoor is de gemiddelde reistijd er langer dan gewenst. Dat is op zich vervelend en kostbaar genoeg, maar wellicht nog schadelijker is de grote variabiliteit van de reistijd op zulke routes. Als het soms een kwartier en soms een uur duurt om naar het werk te gaan, is men gedwongen om ruim op tijd te vertrekken en veelal te vroeg aan te komen. En als men in de file terecht komt, vinden de meeste mensen de onzekerheid over het wel of niet halen van een belangrijke afspraak erger dan het eigenlijke tijdverlies. Dat de reistijd van dag tot dag flink kan verschillen is duidelijk zichtbaar in figuur 2. Tijdens de middagspits variëren de reistijden van 40 minuten tot maar liefst twee uur.

Een reistijdvoorspeller vermindert niet de gemiddelde reistijd, maar kan wel een groot deel van de onzekerheid elimineren. Er blijft natuurlijk nog wel enige onzekerheid over, want het is niet mogelijk om tot op de minuut precies te voorspellen hoe lang het zal duren om van Amsterdam naar Breda te rijden. Hoeveel data, rekenkracht of slimme trucs we ook hebben, we kunnen niet daadwerkelijk in de toekomst kijken. Het blijkt echter, dat als de reistijd op een zekere route een spreiding (standaard deviatie) van, zeg, 20 minuten heeft, onze voorspeller die kan terugbrengen tot 8 minuten. De winst van 12 minuten is zeker zo nuttig als een daadwerkelijke vermindering van de reistijd.

Het probleem

Zij $v(i,d,t)$ de snelheid gemeten bij lus i op dag d en tijd t. We willen de reistijd voorspellen van een reis langs lussen $i = 1,\ldots,I$ als we vertrekken op een zeker tijdstip in de toekomst. In figuur 1 zien we een voorbeeld van het snelheidsveld v voor één dag. Het is zeker niet duidelijk hoe we alle informatie die we hebben verzameld tot op het huidige moment het beste kunnen gebruiken om een reistijdvoorspeller te definiëren. We hebben echter een compressie van deze informatie gevonden, die voor ons doeleinde bijzonder effectief is (zie [3]en [4]).

Merk op dat we met behulp van het snelheidsveld v de reistijd $T(d,t)$ kunnen berekenen bij vertrek op dag d, tijd t. Deze reistijd kunnen we ons voorstellen als behorend bij een pad door het veld v. We merken op dat we informatie van *na* het tijdstip t nodig hebben om de berekening uit te voeren. Met behulp van de informatie die we ter beschikking hebben *op* tijd t kunnen we wel de zogeheten instantane reistijd $T^*(d,t)$ uitrekenen.

$$T^*(d,t) = \sum_{i=1}^{I-1} \frac{2d_i}{v(i,d,t) + v(i+1,d,t)},$$

waarbij d_i de afstand tussen de lussen i en $i+1$ is. De instantane reistijd zou daadwerkelijk worden gerealiseerd als de snelheid na tijd t ongewijzigd zou blijven, totdat de reis voltooid is.

Als we de reistijd $T(d,t)$ hebben berekend voor een aantal dagen $d = 1,2,\ldots,n$ in het verleden, dan kunnen we ook het historisch gemiddelde berekenen

$$\bar{T}(t) = \frac{1}{n}\sum_{d=1}^{n} T(d,t).$$

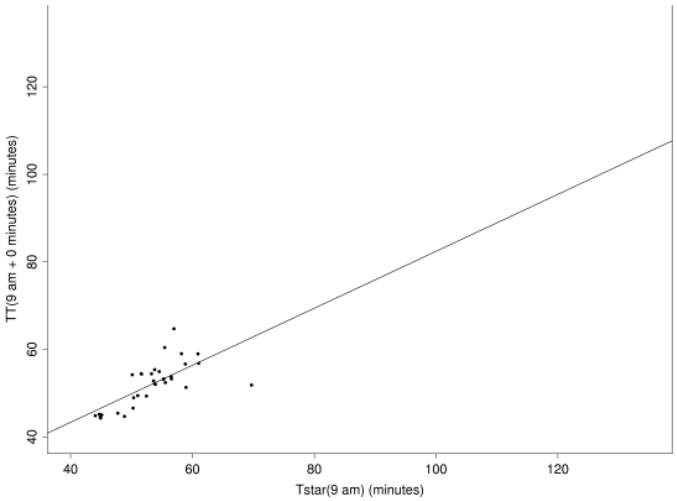
Ons doel is de reistijd $T(d,t+\delta)$, $\delta \geq 0$, te voorspellen met behulp van alle gegevens die we hebben op dag d en tijd t. Hier is δ de tijd tot vertrek, en zoals we al opmerkten is ons probleem ook voor $\delta = 0$ niet triviaal. Twee naïeve voorspellers liggen voor de hand, en worden ook in de praktijk vaak gebruikt. Het historisch gemiddelde $\bar{T}(t+\delta)$ en de instantane reistijd op tijd t, $T^*(d,t)$. We verwachten — en dat blijkt ook zo te zijn — dat $\bar{T}(t+\delta)$ het het beste zal doen voor grote δ, en $T^*(d,t)$ voor kleine δ. Onze nieuwe voorspeller is een gewogen gemiddelde van deze twee naïeve voorspellers, en doet het beter dan beide voor alle δ.

Lineaire regressie

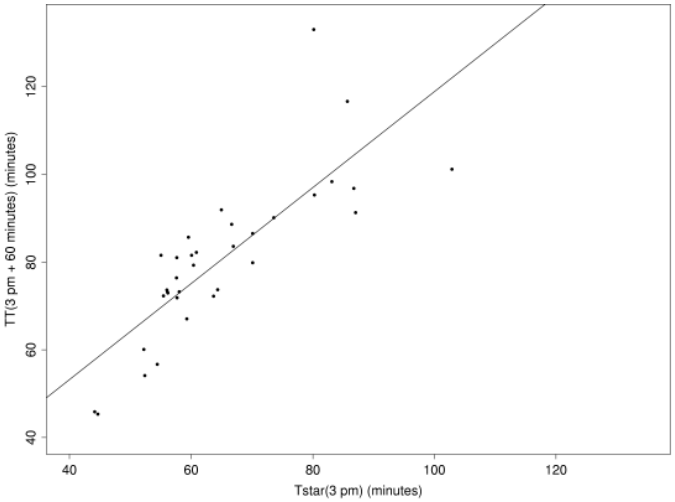
Bij de bestudering van grote hoeveelheden snelwegdata, is ons een empirisch feit opgevallen, namelijk dat er een lineaire relatie bestaat tussen $T^*(d,t)$ en $T(d,t+\delta)$. In figuren 3 en 5 zetten we $T^*(d,t)$ uit tegen $T(d,t+\delta)$ voor onze data van Interstate 10 East. Merk op dat de expliciete relatie varieert met de keuze van t en δ, maar dat de lineairiteit gehandhaafd blijft. Met deze observatie in gedachten, stellen we het volgende model op

$$T(d,t+\delta) = \alpha(t,\delta) + \beta(t,\delta)T^*(d,t) + \varepsilon. \qquad (1)$$

waarbij ε een stochastische grootheid is met verwachting nul die toevallige veranderingen en meetfouten modelleert. Merk op dat



Figuur 3 T^*(9 uur) versus T(9 uur). De regressie lijn snijdt de y-as in het punt α=17,3 en heeft helling β=0,65.



Figuur 5 T^*(15 uur) versus T(16 uur). De regressie lijn snijdt de y-as in het punt α=9,5 en heeft helling β=1,1.

de parameters α en β mogen variëren met t en δ. Lineaire modellen met veranderende parameters worden besproken in [1].

We kunnen $\alpha(t,\delta)$ en $\beta(t,\delta)$ bepalen door de methode van de kleinste kwadraten te gebruiken. Dat wil zeggen, we kiezen $\alpha(t,\delta)$ en $\beta(t,\delta)$ zó dat de kwadratische fout genomen over de historische data zo klein mogelijk is. Met andere woorden, we minimaliseren

$$\sum_{d=1}^{n} (T(d,t+\delta) - \alpha(t,\delta) - \beta(t,\delta)T^*(d,t))^2$$

over α en β. We duiden de optimale parameter waarden aan met $\hat{\alpha}(t,\delta)$ en $\hat{\beta}(t,\delta)$ en definiëren onze voorspeller als

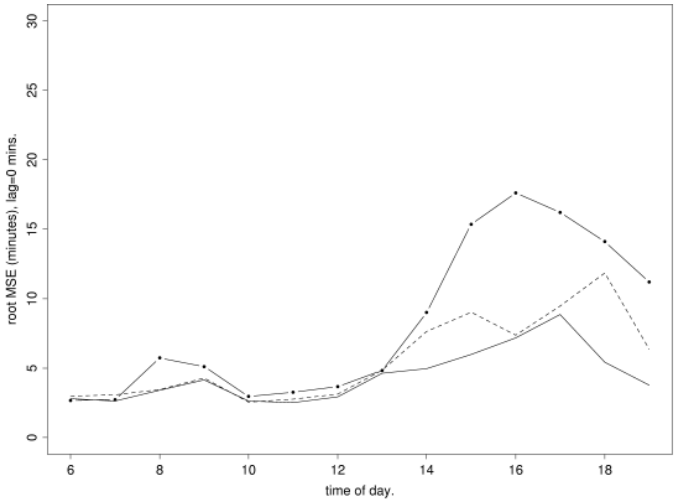
$$\hat{T}(d,t+\delta) = \hat{\alpha}(t,\delta) + \hat{\beta}(t,\delta)T^*(d,t). \qquad (2)$$

Als we $\alpha(t,\delta)$ herschrijven als $\alpha'(t,\delta)\bar{T}(t+\delta)$, dan zien we dat (1) en (2) de toekomstige reistijd uitdrukken als een lineaire combinatie van de twee naïeve voorspellers $\bar{T}(t+\delta)$ en $T^*(d,t)$. We kunnen onze nieuwe voorspeller dus interpreteren als de beste lineaire combinatie van de twee naïeve voorspellers. We mogen verwachten dat onze voorspeller het beter doet dan beide en dat blijkt in de praktijk ook zo te zijn.

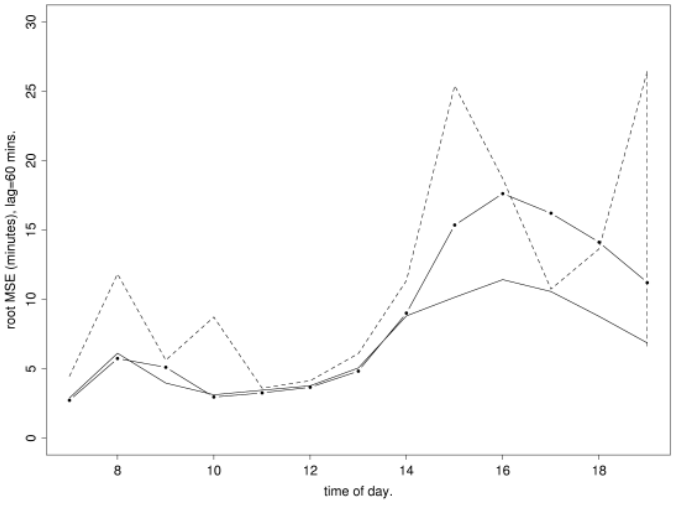
Voorbeeld

We hebben onze voorspeller uitgeprobeerd op onze data van 116 lusdetectoren op een traject van 80 km in Los Angeles. Lusmetingen werden gedaan van 5 uur 's ochtends tot 9 uur 's avonds gedurende 34 weekdagen tussen 16 juni en 8 september 2000. De metingen werden geaggregeerd over intervallen van 5 minuten. Het snelheidsveld v voor 16 juni zien we in figuur 1. De opvallende horizontale lijnen zijn slecht werkende lussen. Het automatisch identificeren en corrigeren van slechte data is een belangrijk probleem, maar we laten dat hier voor wat het is. We hebben de slechte metingen 'met de hand' verwijderd en vervangen door middel van lineaire interpolatie. Vervolgens hebben we de reistijden $T(d,t)$ voor alle dagen berekend, en deze zien we in figuur 2.

We vergelijken de nauwkeurigheid van onze nieuwe voorspel-



Figuur 4 Geschatte wortel van de kwadratische fout voor δ= 0 minuten. Historisch gemiddelde $\bar{T}$ (– · –), instantane reistijd T^* (- - -) en onze voorspeller $\hat{T}$ (—).



Figuur 6 Geschatte wortel van de kwadratische fout voor δ= 60 minuten. Historisch gemiddelde $\bar{T}$ (– · –), instantane reistijd T^* (- - -) en onze voorspeller $\hat{T}$ (—).

ler $\hat{T}(d, t+\delta)$ met de twee naïeve voorspellers $\bar{T}(t+\delta)$ en $T^*(d,t)$ voor verschillende keuzes van t (5 uur, 6 uur,. . ., 20 uur) en δ (0 minuten, 60 minuten). We hebben de wortel van de kwadratische fout geschat door steeds één dag weg te laten, de voorspeller voor die dag te bepalen op basis van de andere dagen, en de kwadratische fouten te middelen. De wortel van de kwadratische fout (root mean squared error, ofwel RMSE) is

$$RMSE(t,\delta) = \left(\frac{1}{34} \sum_{d=1}^{34} (T(d,t+\delta) - \hat{T}(d,t+\delta))^2 \right)^{1/2}$$

De resultaten voor $\delta = 0$ en $\delta = 60$ minuten zijn weergegeven in figuren 4 en 6. Allereerst merken we op dat de instantane reistijd T^* een redelijk goede voorspeller is voor kleine δ, maar zeer slecht voor grote δ. Het historisch gemiddelde is slechter dan T^* voor kleine δ, maar beter voor grote δ. Het belangrijkste resultaat is dat onze nieuwe voorspeller het beter doet dan beide. De RMSE van onze voorspeller blijft onder de 10 minuten, zelfs als we een uur van tevoren voorspellen (δ =60 minuten).

De RMSE van het historisch gemiddelde is een schatting voor de standaard deviatie van de reistijd. We zien in figuur 4 hoe onze voorspeller deze in de middagspits terugbrengt van 20 minuten naar 8.

Opmerkingen

Reistijdvoorspellen is een levendig onderzoeksgebied, en onze voorspeller is zeker niet de enige. Allerlei methoden, bijvoorbeeld gebaseerd op tijdreeks analyse, of op het principe van nabije buren (nearest neighbours), zijn eerder al geprobeerd. Een interessante aanpak die onlangs in Delft is ontwikkeld, maakt gebruik van een neuraal netwerk (zie [2]).

Hèt grote voordeel van onze methode — volgens ons — is zijn eenvoud. De berekening van de optimale parameter waarden $\hat{\alpha}(t,\delta)$ en $\hat{\beta}(t,\delta)$, voor alle t en δ, kost enig gereken, maar dit kan 'off-line' gedaan worden. Om nu op dag d, tijd t een reistijd te voorspellen hoeven we alleen de instantane reistijd $T^*(d,t)$ te berekenen en uit te vermenigvuldigen met de parameters $\hat{\alpha}(t,\delta)$ en $\hat{\beta}(t,\delta)$ voor de gevraagde δ (de tijd tot vertrek).

We hebben onze voorspeller geïmplementeerd voor het netwerk van snelwegen dat Los Angeles doorkruist. We hebben zo'n veertig vertrekpunten en bestemmingen gekozen en voor elk paar de tien kortste routes berekend. Een gebruiker kan via het internet een reistijdvoorspelling en bijbehorende beste route opvragen voor iedere vertrektijd in de toekomst. Als δ groot wordt, convergeert onze voorspeller vanzelf naar het historisch gemiddelde, en telt de huidige verkeerssituatie dus niet meer mee.

Referenties

1 T. Hastie and R. Tibshirani (1993), 'Varying coefficient models', *Journal of the Royal Statistical Society Series B*, **55(4)** p. 757–796.

2 H. van Lint, S.P. Hoogendoorn, H.J. van Zuylen (2002), 'State space neural networks for freeway travel time prediction', *ICANN*, p. 1043–1048.

3 J. Rice and E.W. van Zwet (2001), 'A simple and effective method for predicting travel times on freeways', aangeboden aan: *IEEE Transactions on Intelligent Transportation Systems*.

4 X. Zhang and J. Rice (2001), 'Short-term travel time prediction using a time-varying coefficient lineair model', te verschijnen in: *Transportation Research C*.