Nomograms with two or three straight scales

Dylan Holmes

6 June 2022

Abstract

Nomograms have three scales, each of which may be straight or curved. I show that any nomogram with three straight scales is equivalent to a nomogram with three *parallel* scales $X(u)+Y(v)+W(w)=0$. This characterizes the nomograms made of straight scales.

Any nomogram with *two or three* straight scales can be converted to the functional form $F_1(u)+F_2(v)F_3(w)+F_4(w)=0$ and conversely.

Saint Robert's criterion provides a test for determining whether a function can be represented with three parallel scales and a recipe for constructing the scales if so. I show how to extend the criterion to nomograms with *two or three* straight scales. Thus, if a function can be represented with two or three straight scales, this recipe will produce a nomogram for it.

Three straight scales This is a short proof that if a nomogram is made of three straight-line scales in any orientation, it is equivalent to a nomogram with three *parallel* straight line scales. In other words, you don't get any additional expressive power by allowing straight-line scales to lie in non-parallel orientations. This result is useful because it characterizes the kinds of nomograms you can make with three straight scales.

Proof. Suppose a nomogram is given by three straight scales along distinct lines U, V and W. By using a homography, we may assume without loss of generality that U and V are parallel. If W is already parallel to these two, we are done. Otherwise, W crosses U and V once each; we can use a shear transform to ensure that W is perpendicular to U and V. But this is now a divider-type[1] nomogram which we also know can be represented by three straight scales. $\square$

[1] Also called an N-type or Z-type nomogram. I use the term divider-type because the nomogram is useful for calculating divisions, and because the perpendicular scale visually divides the other two.

Divider-type lemma We made use of the fact that divider-type nomograms can be represented in parallel scale form. Here's the proof.

Proof. The simple divider-type nomogram $F(u,v,w) = \det \begin{bmatrix} 0 & u & 1 \\ 1 & v & 1 \\ w & 0 & 1 \end{bmatrix} = 0$ can be separated into parallel scale form as $\log(u) - \log(v) + \log\left(\frac{w-1}{w}\right) = 0$, which you can confirm using the criterion of Saint Robert. This result holds even if we replace u v and w with smooth functions of those respective variables, and hence (through a suitable homography) it holds for any divider-type nomogram. $\square$

One curved scale By similar reasoning, we can show that any nomogram with two straight scales and one curved scale can be put in the form $F(u,v,w) = F_1(u) + F_2(v)F_3(w) + F_4(w)$, and conversely.

Proof. If a nomogram has one curved scale and two straight scales, then by homography you can map the straight scales onto the vertical lines $x = 0$ and $x = 1$. You end up with a nomogram of the form $F(u,v,w) = \det \begin{bmatrix} u & 0 & 1 \\ v & 1 & 1 \\ f_1(w) & f_2(w) & 1 \end{bmatrix} = 0$. Note the curved W scale. By expanding the determinant, you can see that $F(u,v,w) = u(f_1 - 1) - vf_1 + f_2$.

There's a tradeoff between how nice the determinant looks and how nice the functional form looks. You can swap out the variables f_1 and f_2 with the pair f_3 and f_4, defined by

$$f_3 = -\frac{f_1}{f_1 - 1}, \qquad f_4 = \frac{f_2}{f_1 - 1}$$

This is a reversible transformation whose inverse is:

$$f_1 = \frac{f_3}{1 + f_3}, \qquad f_2 = -\frac{f_4}{1 + f_3}$$

These variables simplify the form of $F(u,v,w) = u + vf_3 + f_4$ at the expense of a more complicated determinant $F(u,v,w) = \det \begin{bmatrix} u & 0 & 1 \\ v & 1 & 1 \\ \frac{f_3}{1+f_3} & -\frac{f_4}{1+f_3} & 1 \end{bmatrix}$. $\square$

Confirming and constructing straight scales Given an equation $F(u,v,w) = 0$, there is a simple test to determine if F can be represented with straight scales, as a nomogram of the form $X(u) + Y(v) + Z(w) = 0$. You take derivatives to compute the quantity

$$R(u,v,w) \equiv \frac{\partial_1 F}{\partial_2 F}.$$

Then u and v have straight scales just if:

$$\partial_1 \partial_2 \log(R) = 0.$$

This is the criterion of Saint Robert. Furthermore, you can construct the functional form of X, Y and Z by computing:

$$\log(\partial_1 X) = \int du\, \partial_1 \log(R) \qquad \text{solve for } X(u)$$
$$\partial_1 Y = R/\partial_1 X \qquad \text{solve for } Y(v)$$

Finally, to find $Z(w)$, we solve $F(u,v,w) = 0$ for u or v; say, we express $\hat{v}$ as a function of the other two variables. We know that $Z(w) = -X(u) - Y(v) = -X(u) - Y(\hat{v}(u,w))$. The function on the left depends only on w; hence we can eliminate u from the right hand side to obtain a pure expression of w.

I haven't seen it mentioned, but an analogous recipe works to check whether a nomogram can be represented with *two or three* straight scales—a more general case. As we've seen, such a nomogram has the form

$$F(u,v,w) = F_1(u) + F_2(v)F_3(w) + F_4(w) = 0.$$

If we compute Saint Robert's R for this expression, we obtain:

$$R = \frac{\partial_1 F}{\partial_2 F} = \frac{F_1'(u)}{F_2'(v) \cdot F_3(w)}.$$

(F_3 is not a derivative.) Note that when F can be represented with *three* straight scales rather than two, R does not depend on w.

Now, behold, $\log(R)$ is the sum of a function of u, a function of v, and a function of w—it has the form $\log(R) = X + Y + Z$ (!). We can therefore recursively use Saint Robert's recipe on the function $\log(R)$ to solve for $F_1'(u)$, $F_2'(v)$ and $F_3(w)$. Finally, $F_4(w) = F(u,v,w) - F_1(u) - F_2(v)F_3(w)$.

1 Finding the parallel scales

Saint Robert's criterion allows you to determine, by taking derivatives, whether $F(u,v,w) = 0$ can be equivalently written in the separated form $X(u) + Y(v) + Z(w) = 0$, in which case it can be nomogrammed using three parallel scales.

Saint Robert also supplies a recipe for actually constructing the functions X, Y, Z. By taking appropriate integrals, you can build two of the functions $X(u)$ and $Y(v)$ out of $F(u,v,w)$ and its derivatives. In the final step, you use the constraint $F(u,v,w) = 0$ and $X(u) + Y(v) + Z(w) = 0$ to solve algebraically for $Z(w)$. Specifically, you start with $Z = -X(u) - Y(v)$, which is a function of u and v, and look for a way to eliminate all references to u and v in favor of w, using the constraint that $F(u,v,w) = 0$.

This last step usually feels like an algebraic miracle. Can you always solve for $Z(w)$ like this? In this section, I'll put the process on firm ground by showing that indeed you can.

What we need to show You have the original function to be nomogrammed, $F(u,v,w) = 0$. You know it is possible to put it in the equivalent form $X(u) + Y(v) + Z(w) = 0$ (such that $F = 0$ if and only if $X + Y + Z = 0$) but so far you only know two of the functions $X(u)$ and $Y(v)$. You are looking for the last function $Z(w)$, which will require some algebraic manipulation.

First, you can use your functions $X(u)$ and $Y(v)$ along with the constraint $X + Y + Z = 0$ to write Z as a function of u and v: $\tilde{Z}(u,v) = -X(u) - Y(v)$.

Next, somehow you have to algebraically manipulate $\tilde{Z}$ as a function of w instead of as a function of u and v. Formally, the constraint $F(u,v,w) = 0$ implicitly allows you to solve for w as a function of u and v. Call this function $\hat{w}(u,v)$. You're looking for a function $Z(w)$ which allows you to factor $\tilde{Z}$ as $\tilde{Z}(u,v) = Z \circ \hat{w}(u,v)$.

To prove that this is always possible, we need two results: first, a powerful theorem that given two functions $f(x,y)$ and $g(x,y)$, you can find a way to factor f as $f = h \circ g$ if and only if the Jacobian of f and g is zero. Next, to apply this theorem, we need to show that indeed the Jacobian of $\tilde{Z}$ and $\hat{w}$ vanishes.

Theorem Suppose f and g are smooth real-valued functions defined on a neighborhood $\Omega \subseteq \mathbb{R}^2$ and whose partial derivatives never vanish[2]. Then the Jacobian $\partial(f,g)$ vanishes if and only if there exists a function $h : \mathbb{R} \to \mathbb{R}$ such that $f = h \circ g$.

Proof. If $f = h \circ g$, then the chain rule shows that the Jacobian of f and g is zero, establishing one direction of the proof.

For the other direction, suppose we solve the equation $g(u,v) = w^\star$ for v. (Formally, we can evoke the implicit function theorem to obtain the unique solution $\hat{v}(u,w^\star) = v$. The required partial derivative $\partial_2 g$ is never zero, by assumption.)

Then $g(u,\hat{v}(u,w^\star)) = w^\star$ throughout, which shows that this expression doesn't actually depend on u. To put it another way, the derivative $\partial_1 g(u,\hat{v}(u,w^\star))$ is zero everywhere. If we expand out this derivative, we get an expression equal to zero that we can solve for $\partial_1 \hat{v}$:

$$\partial_1 \hat{v}(u,w^\star) = -\frac{\partial_1 g(u,\hat{v}(u,w^\star))}{\partial_2 g(u,\hat{v}(u,w^\star))} \tag{1}$$

Next, we know that $\hat{v}(u,g(u,v)) = v$ for any u and v, which means that we can expand $f(u,v)$ into the unwieldy expression $f(u,\hat{v}(u,g(u,v)))$. Defining, for shorthand,

$$\zeta(u,w^\star) \equiv f(u,\hat{v}(u,w^\star)),$$

we have $f(u,v) = f(u,\hat{v}(u,g(u,v))) = \zeta(u,g(u,v))$ everywhere.

[2]Nomogrammable functions have this nondegeneracy property almost everywhere. It ensures that given any two variables, you can solve uniquely for the third. (Or at least, there's a discrete number of solutions, rather than a whole interval of solutions.)

I claim that the value of $\zeta(u, w^\star)$ does not depend on its first argument, in which case we can replace it with the single function $Z(w) = \zeta(u_0, w)$ and have $f(u,v) = Z(g(u,v))$ as we wanted.

Indeed, we can see that the derivative $\partial_1\zeta$ is zero:

$$
\begin{aligned}
\partial_1\zeta &= Df(u,\hat{v})\cdot\langle 1, \partial_1\hat{v}\rangle \\
&= \partial_1 f(u,\hat{v}) + \partial_2 f(u,\hat{v})\cdot\partial_1\hat{v} \\
&\propto \partial_1 g(u,\hat{v}) + \partial_2 g(u,\hat{v})\cdot\partial_1\hat{v} && \{\text{Jacobian } \partial(f,g)=0\} \\
&= \partial_1 g(u,\hat{v}) + \partial_2 g(u,\hat{v})\cdot - \frac{\partial_1 g(u,\hat{v})}{\partial_2 g(u,\hat{v})} && \{\text{eqn (1)}\} \\
&= 0.
\end{aligned}
$$

Here, we've made use of the fact that if the Jacobian of two functions is zero, then their gradients are proportional: $\partial(f,g) = 0 \Leftrightarrow Df \propto Dg$. Indeed, the Jacobian is the dot product of Df with a vector perpendicular to Dg:

$$\partial(f,g) \equiv \partial_1 f\,\partial_2 g - \partial_2 f\,\partial_1 g = \langle \partial_1 f, \partial_2 f\rangle \cdot \langle \partial_2 g, -\partial_1 f\rangle.$$

Hence the dot product is zero if and only if the gradients Df and Dg are parallel.

In short, we've found a function $Z(w) \equiv \zeta(u_0, w) = f(u_0, \hat{v}(u_0, w))$ such that

$$f(u,v) = \zeta(u, g(u,v)) = \zeta(u_0, g(u,v)) = Z(g(u,v)) = Z\circ g$$

as required.

$\square$