Let’s talk about optimization, and a classical tool is the so-called conjugate. Given a function f:\mathbb{R}^p\to\mathbb{R} its conjugate is function f^{\star}:\mathbb{R}^p\to\mathbb{R} such that f^{\star}(\boldsymbol{y})=\max\_{\boldsymbol{x}}

\lbrace\boldsymbol{x}^\top\boldsymbol{y}-f(\boldsymbol{x})\rbraceso, long story short, f^{\star} (\boldsymbol{y}) is the maximum gap between the linear function \boldsymbol{x}^\top\boldsymbol{y} and f(\boldsymbol{x}).

Just to visualize, consider a simple parabolic function (in dimension 1) f(x)=x^2/2, then f^{\star} (\color{blue}{2}) is the maximum gap between the line x\mapsto\color{blue}{2}x and function f(x).

x = seq(-100,100,length=6001) f = function(x) x^2/2

vf = Vectorize(f)(x)

fstar = function(y) max(y\*x-vf) vfstar = Vectorize(fstar)(x)

We can see it on the figure below.

viz = function(x0=1,YL=NA){

idx=which(abs(x)<=3) par(mfrow=c(1,2)) plot(x[idx],vf[idx],type="l", xlab="",ylab="",col="blue",lwd=2) abline(h=0,col="grey") abline(v=0,col="grey") idx2=which(x0\*x>=vf) polygon(c(x[idx2],rev(x[idx2])),c(vf[idx2],rev(x0\*x[idx2])), col=rgb(0,1,0,.3),border=NA)

abline(a=0,b=x0,col="red") i=which.max(x0\*x-vf)

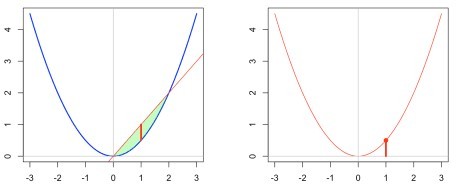
segments(x[i],x0\*x[i],x[i],f(x[i]),lwd=3,col="red") if(is.na(YL)) YL=range(vfstar[idx])

plot(x[idx],vfstar[idx],type="l",xlab="",ylab="",col="red",lwd=1,ylim=YL) abline(h=0,col="grey")

abline(v=0,col="grey") segments(x0,0,x0,fstar(x0),lwd=3,col="red") points(x0,fstar(x0),pch=19,col="red")

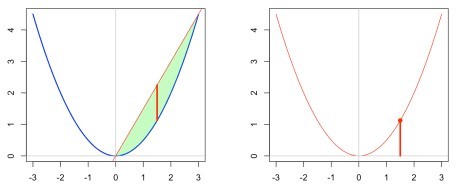
}

viz(1)



or

viz(1.5)



In that case, we can actually compute f^{\star}, since f^{\star}(y)=\max\_{x}\lbrace xy-f(x)\rbrace=\max\_{x}

\lbrace xy-x^2/2\rbraceThe first order condition is here x^{\star}=y and thusf^{\star}(y)=\max\_{x}\lbrace xy- x^2/2\rbrace=\lbrace x^{\star}y-(x^{\star})^2/2\rbrace=\lbrace y^2-y^2/2\rbrace=y^2/2And actually, that can be related to two results. The first one is to observe that f(\boldsymbol{x})=\|\boldsymbol{x}\|\_2^2/2 and in that case f^{\star}(\boldsymbol{y})=\|\boldsymbol{y}\|\_2^2/2 from the following general result : if f(\boldsymbol{x})=\|\boldsymbol{x}\|\_p^p/p with p>1, where \|\cdot\|\_p denotes the standard \ell\_p norm, then f^{\star}(\boldsymbol{y})=\|\boldsymbol{y}\|\_q^q/q where\frac{1}{p}+\frac{1}{q}=1The second one is the conjugate of a quadratic function. More specifically if f(\boldsymbol{x})=\boldsymbol{x}^{\top}\boldsymbol{Q}\ boldsymbol{x}/2 for some definite positive matrix \boldsymbol{Q}, f^{\star}(\boldsymbol{y})=\ boldsymbol{y}^{\top}\boldsymbol{Q}^{-1}\boldsymbol{y}/2. In our case, it was a univariate problem with

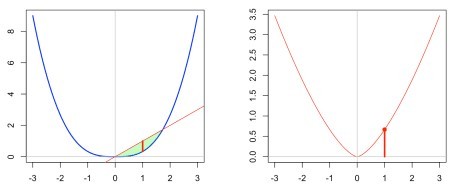
\boldsymbol{Q}=1.

For the conjugate of the \ell\_p norm, we can use the following code to visualize it

p = 3

f = function(x) abs(x)^p/p vf = Vectorize(f)(x)

fstar = function(y) max(y\*x-vf) vfstar = Vectorize(fstar)(x) viz(1.5)

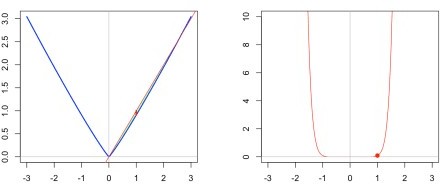


or

p = 1.1

f = function(x) abs(x)^p/p vf = Vectorize(f)(x)

fstar = function(y) max(y\*x-vf) vfstar = Vectorize(fstar)(x) viz(1, YL=c(0,10))



Actually, in that case, we almost visualize that if f(x)=|x| then\displaystyle{f^{\star}\left(y\right)=

{\begin{cases}0,&\left|y\right|\leq 1\\\infty ,&\left|y\right|>1.\end{cases}}}

To conclude, another popular case, f(x)=\exp(x) then{\displaystyle f^{\star}\left(y\right)={\begin{cases}y\log(y)- y,&y>0\\0,&y=0\\\infty ,&y<0.\end{cases}}}[/latex]We can visualize that case below

f = function(x) exp(x) vf = Vectorize(f)(x)

fstar = function(y) max(y\*x-vf) vfstar = Vectorize(fstar)(x) viz(1,YL=c(-3,3))

