Constructing Diversification Constraints

R. Scott McIntire

Oct 5, 2024

1 Overview

Consider the function function, $f : \mathbf{R}^n \mapsto \mathbf{R}$, defined by

$$f(\mathbf{x}) \;=\; \sum_{i=1}^{k} x_{[i]} \tag{1}$$

From this we see that $f(\mathbf{x})$ is the value of the sum of the k^{th} largest values of its input $\mathbf{x}$. We are interested in optimization problems involving a vector $\mathbf{x}$ with the sum of the top values constrained by a given value; that is, such that $f(\mathbf{x}) \leq M$ for some value M. How could we do this? One way is to write down all possible combinations of the k elements of $\mathbf{x}$ and write a constraint that bounds their sum to be less than or equal to M. But the number of constraints that one has to write is $\binom{n}{k}$. This becomes large very quickly. In the next section we seek a way to represent f to reduce the number of constraints.

This constraint often arises in the world of finance where one wants to create (or modify) a portfolio so that the new portfolio has the property that its top k securities (notionally) do not exceed some fixed percentage of its total notional.

2 The Function f as an Optimization Problem

We wish to characterize the value of f on a given vector, $\mathbf{x}$ as the result of an optimization. We claim that for for a given $\mathbf{x}$, $f(\mathbf{x})$ is the value of the following constrained optimization problem:[1]

$$\max_{\mathbf{y}\in\mathbf{Z}_2^n} \quad \mathbf{y}^T\mathbf{x} \tag{2}$$

$$\mathbf{y}^T\mathbf{1} \;=\; k \tag{3}$$

In English this says: "Take the maximum value of all possible sums of k values from $\mathbf{x}$." This is clearly just another way of restating $f(\mathbf{x})$. Although this optimization is succinct, there are still an exponential number of combinations to examine to find the optimal solution in this discrete setting.

claim: The solution to the above is the same as the solution to the *continuous* optimization problem:

$$\max_{\mathbf{y}\in\mathbf{R}^n} \quad \mathbf{y}^T\mathbf{x} \tag{4}$$

$$\mathbf{0} \preceq \mathbf{y} \preceq \mathbf{1} \tag{5}$$

$$\mathbf{y}^T\mathbf{1} = k \tag{6}$$

proof: Since its domain, $\mathbf{R}^n$, includes $\mathbf{Z}_2^n$, the solution to the continuous problem is at least as large as the solution to the original discrete problem. We now show the reverse, by showing that for every optimal solution to the continuous problem, $\mathbf{y}^*$, there is a corresponding $\mathbf{y}'$ that lives in $\mathbf{Z}_2^n$ yielding the same value for the objective.

To this end, let $\mathbf{y}^*$ be a solution to the continuous problem. If all of the values are integral; meaning, take on only the values of 0 or 1 we are done. Otherwise, collect the components of $\mathbf{y}^*$ that are non-integral. By the constraint, (6), there must be more than one component. It must also be the case that amongst these "fractional" components that the corresponding "x" values are the same. Otherwise, If two of the components of $\mathbf{y}$ had differing "x" values, then the "y" component with the smaller "x" value could shift some of its value to the "y" component with the larger "x" value in a way that keeps the constraint, (6). But then we would have a larger objective value that the supposed maximum – contradiction. Therefore, all of the "x" components associated with non-integral "y" values must have the same value. The sum of these non-integral values must sum to an integral value, which by necessity, is less than the number of components in $\mathbf{y}^*$. Suppose this number is j. Take any j of these "y" components. Create a new vector $\mathbf{y}'$ and set its values in the following way: Set the

[1] A bold font is used to denote vectors. We refer to a vector of 0s as $\mathbf{0}$, and likewise we refer to a vector of 1s as $\mathbf{1}$. Finally, the statement, $\mathbf{a} \preceq \mathbf{b}$, is the statement that every element of the vector $\mathbf{a}$ is less than or equal to the corresponding element in the vector $\mathbf{b}$. The statement, $\mathbf{a} \succeq \mathbf{b}$, is similar but with each element of $\mathbf{a}$ greater than or equal each corresponding element of $\mathbf{b}$.

components of $\mathbf{y}'$ corresponding to the previous j components from $\mathbf{y}^*$ to be 1. Set the components of $\mathbf{y}'$ that correspond to the remaining "fractional" components of $\mathbf{y}^*$ to 0. Set the rest of the components of $\mathbf{y}'$ to the corresponding values from $\mathbf{y}^*$. We have now shown that a solution vector to the continuous problem is in the feasible set of the discrete problem. Therefore, the optimal solution to the discrete problem is at least as large as the solution to the continuous problem. Consequently, the two problems are equivalent.

Note that the solution to continuous problem is the same as the associated problem:

$$\min_{\mathbf{y}\in\mathbf{R}^n} \quad -\mathbf{y}^T\mathbf{x} \tag{7}$$
$$\mathbf{0} \preceq \mathbf{y} \preceq \mathbf{1} \tag{8}$$
$$\mathbf{y}^T\mathbf{1} = k \tag{9}$$

But this problem is a *convex* problem.

3 A Dual Description of the Optimization

Since the problem described by equations $(7, 8, 9)$, is a *convex* problem, the value of its solution is the same as the value of its associated *dual* problem.[2]

To form the dual problem we need the Lagrangian, which is:[3]

$$L(\mathbf{y}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \nu; \mathbf{x}) = -\mathbf{y}^T\mathbf{x} - \boldsymbol{\lambda}_1^T\mathbf{y} + \boldsymbol{\lambda}_2^T(\mathbf{y} - \mathbf{1}) + \nu(k - \mathbf{y}^T\mathbf{1}) \tag{10}$$

Set the function g:

$$g(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_1, \nu; \mathbf{x}) = \inf_{\mathbf{y}} L(\mathbf{y}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \nu; \mathbf{x}) \tag{11}$$

Substituting for L this becomes

$$g(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_1, \nu; \mathbf{x}) = \inf_{\mathbf{y}} \left(\mathbf{y}^T(-\mathbf{x} - \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 - \nu\mathbf{1}) - \boldsymbol{\lambda}_2^T\mathbf{1} + \nu\, k\right) \tag{12}$$

[2] Normally one needs to show that a convex problem satisfies Slater's condition in order to claim that it provides the same value as its associated dual. But this is not necessary when dealing with *linear* convex problems. Why, you may ask is it *linear*? The term $-\mathbf{y}^T\mathbf{x}$ is decidedly non-linear. It is linear if we treat $\mathbf{x}$ as a constant (we also use the term parameter).

[3] For the purposes of this optimization, $\mathbf{x}$ is seen as a parameter to the optimization – essentially a constant with respect to the optimization. As such, it is customary to treat such parameters as part of a function but separate them from the "true" list of variables of the problem with a semi-colon.

The dual problem is then

$$\max_{\substack{\boldsymbol{\lambda}_2\succeq\mathbf{0}\\ \boldsymbol{\lambda}_1\succeq\mathbf{0}\\ \nu}} g(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\nu;\mathbf{x}) \tag{13}$$

Which is[4]

$$\max_{\substack{\boldsymbol{\lambda}_2\succeq\mathbf{0}\\ \boldsymbol{\lambda}_1\succeq\mathbf{0}\\ \nu}} -\boldsymbol{\lambda}_2^T\mathbf{1}+\nu\, k \tag{14}$$

$$-\boldsymbol{\lambda}_1+\boldsymbol{\lambda}_2-\nu\mathbf{1} \quad = \quad \mathbf{x} \tag{15}$$

This is equivalent to[5]

$$\max_{\substack{\boldsymbol{\lambda}\succeq\mathbf{0}\\ \nu}} \quad -\boldsymbol{\lambda}^T\mathbf{1}+\nu\, k \tag{16}$$

$$\boldsymbol{\lambda}-\nu\mathbf{1}\succeq\mathbf{x} \tag{17}$$

Or,

$$\max_{\substack{\boldsymbol{\lambda}\succeq\mathbf{0}\\ \nu}} \qquad \nu\, k-\mathbf{1}^T\boldsymbol{\lambda} \tag{18}$$

$$\mathbf{x}\preceq\boldsymbol{\lambda}-\nu\mathbf{1} \tag{19}$$

Since maximizing over ν or $-\nu$ is the same and there are no restrictions on the sign of ν we may replace ν with $-\nu$ in the last equations giving:

$$\max_{\substack{\boldsymbol{\lambda}\succeq\mathbf{0}\\ \nu}} \qquad -\nu\, k-\mathbf{1}^T\boldsymbol{\lambda} \tag{20}$$

$$\mathbf{x}\preceq\boldsymbol{\lambda}+\nu\mathbf{1} \tag{21}$$

But this is the same as:

$$\min_{\boldsymbol{\lambda},\nu} \qquad \nu\, k+\mathbf{1}^T\boldsymbol{\lambda} \tag{22}$$

$$\mathbf{x}\preceq\boldsymbol{\lambda}+\nu\mathbf{1} \tag{23}$$

$$\boldsymbol{\lambda}\succeq\mathbf{0} \tag{24}$$

[4]Note that g is $-\infty$, unless the term that $\mathbf{y}$ is "dotting" is the zero vector. Consequently, the maximum must necessarily occur where the dotting vector is zero.

[5]We can remove n variables by eliminating $\boldsymbol{\lambda}_1$ from the equations while keeping the same number of inequalities. We can do this by realizing that $\boldsymbol{\lambda}_2-\nu\mathbf{1}=\mathbf{x}+\boldsymbol{\lambda}_1$ expresses the same information as: $\boldsymbol{\lambda}_2-\nu\mathbf{1}\succeq\mathbf{x}$. Since there is now only one $\boldsymbol{\lambda}$, we relabel $\boldsymbol{\lambda}_2$ as $\boldsymbol{\lambda}$.

4 Characterization of f as O(n) Linear Constraints

If one wishes to bound the top k elements of the vector $\mathbf{x}$ by M in an optimization problem; that is, to bound $f(\mathbf{x})$ above by M, one needs to add two new variables, $\boldsymbol{\lambda}$ and ν along with the following three constraints:

$$\begin{aligned} \nu\, k + \boldsymbol{\lambda}^T \mathbf{1} &\leq M && (25) \\ \mathbf{x} &\preceq \boldsymbol{\lambda} + \nu \mathbf{1} && (26) \\ \boldsymbol{\lambda} &\succeq \mathbf{0} && (27) \end{aligned}$$

Why? Because the expression $\nu\, k + \boldsymbol{\lambda}^T \mathbf{1}$ with the constraints (26) and (27) applied is *always* an upper bound to $f(\mathbf{x})$. Consequently, we have the inequality: $f(\mathbf{x}) \leq \nu\, k + \boldsymbol{\lambda}^T \mathbf{1} \leq M$. The only concern remaining is that there may be a gap between the values of $\nu\, k + \boldsymbol{\lambda}^T \mathbf{1}$ and $f(\mathbf{x})$ – making (25) too restrictive. But this is not the case as the minimum over all $\boldsymbol{\lambda}$ and ν (subject to (26) and (27)) is $f(\mathbf{x})$. Another way of saying this is that the parameters, $\boldsymbol{\lambda}$ and ν provide enough freedom to $\mathbf{x}$ so that $f(\mathbf{x})$ can be as close to M as one would like.

Therefore, in order to avoid a combinatorial explosion of inequality constraints, one need only add $(n+1)$ variables, $(\boldsymbol{\lambda}, \nu)$, to an optimization problem to provide diversification constraints on a vector of length n. The number of new inequality constraints added becomes $(2n + 1)$.

One might ask why we couldn't do exactly the same procedure with the "simpler" characterization we had with equations (7), (8), and (9), rather than going through so much effort with a dual problem. The same number of variables and the same number of constraints are required.

We could do that; however, note that once we take the characterization of a bound on a "given" $\mathbf{x}$ and place it into an optimization problem, that "given" $\mathbf{x}$ is no longer a constant. It, itself, is part of the optimization. And now, one of the inequalities that we would add would be related to (7); that is, the term: $-\mathbf{y}^T \mathbf{x}$. But this term involves the (now) *variable*, $\mathbf{x}$, and the new *variable*, $\mathbf{y}$, in a decidedly *non-linear* way. In fact, the term is not even *convex*.

In contrast to this "simpler" characterization, we have shown that by examining an associated dual problem, $f(\mathbf{x})$ can be bounded if we add new variables and new *linear* constraints. What's more, the numbers of variables and constraints are linear in the length of the vector $\mathbf{x}$.