Fucking Arb Problem

Let's start by writing down the original problem, for reference:

$$\text{maximize: } U(\Psi) \tag{1}$$

$$\text{subject to: } \Psi = \sum_{i=1}^{m} A_i(\Lambda_i - \Delta_i) \tag{2}$$

$$\varphi_i(R_i + \gamma_i\Delta_i - \Lambda_i) = \varphi_i(R_i), \ \ (i = 1, ..., m) \tag{3}$$

$$\Delta_i \geq 0, \ \ \Lambda_i \geq 0, \ \ (i = 1, ..., m) \tag{4}$$

Ok, let's dramatically simplify the presentation of this problem. First, the variables Λ_i and Δ_i represent the vectors of tendered/received tokens on the i^{th} market, written with respect to the local (market specific) index (where the matrix A_i transforms them into the global indexing scheme). So, I hate four things about this notation, and I'm going to fix them.

- First, I keep forgetting which is tendered and which is received, so I'm going to use the letters (t, r) instead of (Λ, Δ).
- Second, I keep forgetting that they are vectors, so I'm going to use vector arrows: $(\vec{t}, \vec{r})$
- Third, fuck the local indexing. Let's just assume they are already written in terms of the global indexing scheme. In other words, I'll take *as definition* the vectors $\vec{t_i}$ and $\vec{r_i}$ to be already transformed by A_i:

$$\vec{t_i} := A_i\Lambda_i \tag{5}$$

$$\vec{r_i} := A_i\Delta_i \tag{6}$$

- Fourth, I want to use the index i for something else (my *super-vector* that we'll see later). So I'm going to use the index m to keep track of the separate markets, with a total of M markets. Therefore, instead of having $i = 1, .., m$, I'm using $m = 1, ..., M$.

With all of these changes in mind, then the original problem (1)-(4) can be rephrased as the following:

$$\text{maximize: } U(\vec{\Psi}) \tag{7}$$

$$\text{subject to: } \vec{\Psi} = \sum_{m=1}^{M} (\vec{t}_m - \vec{r}_m) \tag{8}$$

$$\varphi_m(\vec{R}_m + \gamma_m\vec{t}_m - \vec{r}_m) = \varphi_m(\vec{R}_m), \ \ (m = 1, ..., M) \tag{9}$$

$$\vec{t}_m \geq 0, \ \ \vec{r}_m \geq 0, \ \ (m = 1, ..., M) \tag{10}$$

But there's one more change I want to make. Since we already decided that our utility function will be a simple weighted sum, i.e. $\vec{u}^T \cdot \vec{\Psi}$ for some utility vector $\vec{u}$, then let's just combine (7) and (8) into one line (no need for the Ψ middleman variable):

$$\text{maximize: } \vec{u}^T \cdot \left(\sum_{m=1}^{M} (\vec{t}_m - \vec{r}_m) \right) \tag{11}$$

$$\text{subject to: } \varphi_m(\vec{R}_m + \gamma_m\vec{t}_m - \vec{r}_m) = \varphi_m(\vec{R}_m), \ \ (m = 1, ..., M) \tag{12}$$

$$\vec{t}_m \geq 0, \ \ \vec{r}_m \geq 0, \ \ (m = 1, ..., M) \tag{13}$$

And lastly, just to be clear,

- the quantities $\vec{R}_m$, γ_m, and $\vec{u}$ are *given in advance*
- the quantities $\vec{t}_m$ and $\vec{r}_m$ are the *dynamic variables* with which we want to maximize

Ok but there's actually one more thing I want to do to rephrase the problem, on the next page.

Let's define the variable $\vec{x}$ consisting of the *stack* of all of our $\vec{t}_m$'s and $\vec{r}_m$'s.

$$\vec{x} := \begin{bmatrix} [\vec{t}_1] \\ [\vec{r}_1] \\ [\vec{t}_2] \\ [\vec{r}_2] \\ \vdots \\ [\vec{t}_M] \\ [\vec{r}_M] \end{bmatrix} \tag{14}$$

Furthermore, we note that the objective function given in (11) can be written as

$$\begin{aligned}
\vec{u}^T\cdot \left(\sum_{m=1}^{M} (\vec{t}_m - \vec{r}_m)\right) &= \vec{u}^T\cdot \Big((\vec{t}_1 - \vec{r}_1) \ + \ (\vec{t}_2 - \vec{r}_2) \ + \ ... \ + \ (\vec{t}_M - \vec{r}_M)\Big) \\
&= \vec{u}^T\cdot \vec{t}_1 \ - \ \vec{u}^T\cdot \vec{r}_1 \ + \ \vec{u}^T\cdot \vec{t}_2 \ - \ \vec{u}^T\cdot \vec{r}_2 \ + \ ... \ + \ \vec{u}^T\cdot \vec{t}_M \ - \ \vec{u}^T\cdot \vec{r}_M \\
&= \Big[\ [\ \vec{u}^T \ ] \ [\ -\vec{u}^T \ ] \ [\ \vec{u}^T \ ] \ [\ -\vec{u}^T \ ] \ ... \ [\ \vec{u}^T \ ] \ [\ -\vec{u}^T \ ] \ \Big] \cdot \begin{bmatrix} [\vec{t}_1] \\ [\vec{r}_1] \\ [\vec{t}_2] \\ [\vec{r}_2] \\ \vdots \\ [\vec{t}_M] \\ [\vec{r}_M] \end{bmatrix} &(15) \\
&= \vec{w}^T\cdot \vec{x} &(16)
\end{aligned}$$

where I've defined the vector $\vec{w}$ as a bunch of alternating copies of $\vec{u}$:

$$\vec{w}^T := \Big[\ [\ \vec{u}^T \ ] \ [\ -\vec{u}^T \ ] \ [\ \vec{u}^T \ ] \ [\ -\vec{u}^T \ ] \ ... \ [\ \vec{u}^T \ ] \ [\ -\vec{u}^T \ ] \ \Big] \tag{17}$$

Finally, let's define the functions h_m that will encode the market maker formulas expressed in (12):

$$h_m\Big(\vec{x}\Big) := \varphi_m\Big(\vec{R}_m + \gamma_m \vec{t}_m - \vec{r}_m\Big) - \varphi_m\Big(\vec{R}_m\Big) \tag{18}$$

Note that it is legit to say that h_m depends on $\vec{x}$, even though it only depends on the *particular chunck* of $\vec{x}$ given by $\vec{t}_m$ and $\vec{r}_m$. With all of these changes to our notation, we can once again rephrase the problem. Statements (11)-(13) now become

$$\begin{aligned}
\text{maximize: } & f(\vec{x}) = \vec{w}^{\,T}\cdot \vec{x} &(19) \\
\text{subject to: } & h_m\Big(\vec{x}\Big) = 0 \ \ (m = 1, ..., M) &(20) \\
& \vec{x} \geq 0 &(21)
\end{aligned}$$

Ok, one more thing... not every token is traded on every market, and so some of the components of $\vec{x}$ must be fixed at zero. For example, suppose that token k (in the global indexing) is *not* traded on market m. Then we must have $\big(\vec{t}_m\big)_k = \big(\vec{r}_m\big)_k = 0$, always and forever. Since the vector $\vec{x}$ is just a stack of $\vec{t}$'s and $\vec{r}$'s, then this would result in $\big(\vec{x}\big)_{i_0} = \big(\vec{x}\big)_{j_0} = 0$, for some particular i_0 and j_0.

In light of this, let the index i range over all of $\vec{x}$, and then let us introduce the following vocabulary: define an *active index* i to be one that is *not fixed* to be zero. Similarly, we define an *inactive index* i to be one that *is fixed* to be zero. Rather than imposing additional constraints such as $\vec{x}_i = 0$ for all inactive i, we should really just remove all of the inactive variables altogether from the whole problem (19)-(21), since these inactive variables are static and immaterial. The process of removing the inactive variables out of $\vec{x}$ will be made explicit when I work out an example later on.

Now, on to the fucking KKT conditions.

Supposing now that we've removed all the inactive variables from our vector $\vec{x}$ (and consequently remove them from $\vec{w}$ as well, so that the dimensions match), then let's define the following functions:

$$g_i(\vec{x}) = -x_i \tag{22}$$

i.e. it's just the negative i^{th} component of $\vec{x}$. With this, then we can rephrase our problem yet again:

$$\text{maximize: } f(\vec{x}) = \vec{w}^{\,T}\!\cdot \vec{x} \tag{23}$$

$$\text{subject to: } h_m\Big(\vec{x}\Big) = 0 \;\; (m = 1, ..., M) \tag{24}$$

$$g_i(\vec{x}) \leq 0 \;\; (i = 1, ..., \text{length of } \vec{x}) \tag{25}$$

As far as I can tell, this setup is perfectly applicable to premises of the KKT theorem.

In particular, the internet says that if $\vec{x}^*$ is a solution to the problem (23)-(25), then it can be characterized as follows. There exist values μ_i $(i = 1, ..., \text{length of } \vec{x})$ and λ_m $(m = 1, ..., M)$ such that the following conditions are met:

$\boxed{\textbf{Primal Feasibility}}$

$$h_m\big(\vec{x}^*\big) = 0 \;\; (m = 1, ..., M) \tag{26}$$

$$g_i\big(\vec{x}^*\big) \leq 0 \;\; (i = 1, ..., \text{length of } \vec{x}) \tag{27}$$

$\boxed{\textbf{Dual Feasibility}}$

$$\nabla L(\vec{x}^*) = \nabla \left(f(\vec{x}) - \sum_i \mu_i g_i(\vec{x}) - \sum_m \lambda_m h_m(\vec{x}) \right)\Bigg|_{\vec{x}^*} = 0 \tag{28}$$

$$\mu_i \geq 0 \tag{29}$$

$\boxed{\textbf{Complementary Slackness}}$

$$\mu_i \; g_i\big(\vec{x}\big) = 0 \;\; (\text{for each } i) \tag{30}$$

In other words, we can interpret these conditions as follows:

- **Primal Feasibility:** If you think you've found a solution $\vec{x}^*$, it better at least satisfy the original constraints (24) and (25).
- **Dual Feasibility:** More interestingly, the solution $\vec{x}^*$ will be a critical point for the Lagrangian function defined by $L(\vec{x}) := f(\vec{x}) - \sum_i \mu_i g_i(\vec{x}) - \sum_m \lambda_m h_m(\vec{x})$. Also, the values of μ_i will be non-negative.
- **Complementary Slackness:** Of the inequality constraints in (25) ranging over i, those who are honest inequalities (and not just equal to zero) require the corresponding μ_i to be zero.

Now, notice that for us, the Lagrangian takes a pretty simple form. We already know that our objective function is $f(\vec{x}) = \vec{w}^{\,T}\!\cdot \vec{x}$. But, if we arrange our collection of μ_i values in a vector $\vec{\mu}$, then in light of (22), the expression $\sum_i \mu_i g_i(\vec{x})$ can be written as

$$\sum\nolimits_i \mu_i g_i(\vec{x}) = -\vec{\mu}^{\,T}\!\cdot \vec{x} \tag{31}$$

Therefore, the Lagrangian takes the somewhat simpler form of

$$L(\vec{x}) = \Big(\vec{w} + \vec{\mu}\Big)^T\!\cdot \vec{x} - \sum\nolimits_m \lambda_m h_m(\vec{x}) \tag{32}$$

Ok ok ok. Now let's do an explicit example, to see where it all goes wrong...

Simple Example

Let's suppose we have 3 global tokens $\{A, B, C\}$, and 3 total markets $\{1, 2, 3\}$. Let's suppose that market 1 only trades $\{A, B\}$, while market 2 only trades $\{B, C\}$, and market 3 only trades $\{A, C\}$. Let's summarize this in the following table, and I will explicitly write down which variables x_i appear in what tendered and received vectors:

Market:	1	2	3
Tokens Traded:	$\{A, B\}$	$\{B, C\}$	$\{A, C\}$
Tendered Vector:	$\vec{t}_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 = 0 \end{bmatrix}$	$\vec{t}_2 = \begin{bmatrix} x_7 = 0 \\ x_8 \\ x_9 \end{bmatrix}$	$\vec{t}_3 = \begin{bmatrix} x_{13} \\ x_{14} = 0 \\ x_{15} \end{bmatrix}$
Received Vector:	$\vec{r}_1 = \begin{bmatrix} x_4 \\ x_5 \\ x_6 = 0 \end{bmatrix}$	$\vec{r}_2 = \begin{bmatrix} x_{10} = 0 \\ x_{11} \\ x_{12} \end{bmatrix}$	$\vec{r}_3 = \begin{bmatrix} x_{16} \\ x_{17} = 0 \\ x_{18} \end{bmatrix}$

The red quantities represent the inactive variables. If we remove the inactive variables, and then stack up our tendered and received vectors, we get our vector $\vec{x}$:

$$\vec{x} = \begin{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \begin{bmatrix} x_4 \\ x_5 \end{bmatrix} \\ \begin{bmatrix} x_8 \\ x_9 \end{bmatrix} \\ \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} \\ \begin{bmatrix} x_{13} \\ x_{15} \end{bmatrix} \\ \begin{bmatrix} x_{16} \\ x_{18} \end{bmatrix} \end{bmatrix} \tag{33}$$

(Notice the inactive x_i's missing from 33). Similarly, for the utility vector $\vec{u}^{\,T} = [u_A, u_B, u_C]$, we have

$$\vec{w}^{\,T} = \Big[[u_A, u_B], [-u_A, -u_B], [u_B, u_C], [-u_B, -u_C], [u_A, u_C], [-u_A, -u_C] \Big] \tag{34}$$

For our market maker functions h_m, let's just assume that they are simple constant product market makers, for this example. This means that the function φ_m in (18) is just the product of all of its inputs: $\varphi\big([x, y]\big) = xy$. For the reserve vectors

$$\vec{R}_1 = \begin{bmatrix} R_{1A} \\ R_{1B} \\ 0 \end{bmatrix}, \quad \vec{R}_2 = \begin{bmatrix} 0 \\ R_{2B} \\ R_{2C} \end{bmatrix}, \quad \vec{R}_3 = \begin{bmatrix} R_{3A} \\ 0 \\ R_{3C} \end{bmatrix}, \tag{35}$$

our constant product functions become

$$h_1(\vec{x}) = \Big(R_{1A} + \gamma_1\, x_1 - x_4\Big)\Big(R_{1B} + \gamma_1\, x_2 - x_5\Big) - \Big(R_{1A}\Big)\Big(R_{1B}\Big) \tag{36}$$

$$h_2(\vec{x}) = \Big(R_{2B} + \gamma_2\, x_8 - x_{11}\Big)\Big(R_{2C} + \gamma_2\, x_9 - x_{12}\Big) - \Big(R_{2B}\Big)\Big(R_{2C}\Big) \tag{37}$$

$$h_3(\vec{x}) = \Big(R_{3A} + \gamma_3\, x_{13} - x_{16}\Big)\Big(R_{3C} + \gamma_3\, x_{15} - x_{18}\Big) - \Big(R_{3A}\Big)\Big(R_{3C}\Big) \tag{38}$$

Again, these equation (36)-(38) are derived starting from the definition of h_m in (18), using the constant product formula $\varphi\big([x, y]\big) = xy$, and converting all $\vec{t}$ and $\vec{r}$ into x_i's using the table above. Next, we set up the Lagrangian.

Ok, so we take the expression for the Lagrangian we found in (32)

$$L(\vec{x}) = \Big(\vec{w} + \vec{\mu}\Big)^T \cdot \vec{x} - \sum\nolimits_m \lambda_m h_m(\vec{x}) \tag{39}$$

and take the gradient:

$$\begin{aligned} \nabla L(\vec{x}) &= \nabla \left[\Big(\vec{w} + \vec{\mu}\Big)^T \cdot \vec{x} \right] - \nabla \left[\sum\nolimits_m \lambda_m h_m(\vec{x}) \right] \\ &= \Big(\vec{w} + \vec{\mu}\Big) - \nabla \left[\sum\nolimits_m \lambda_m h_m(\vec{x}) \right] \end{aligned} \tag{40}$$

where the second line above follows from the fact that $\nabla\big(\vec{v}^{\,T} \cdot \vec{x}\big) = \vec{v}$. So, if we are trying to solve the system $\nabla L(\vec{x}) = 0$, we set (40) equal to zero:

$$\Big(\vec{w} + \vec{\mu}\Big) = \nabla \left[\sum\nolimits_m \lambda_m h_m(\vec{x}) \right] \tag{41}$$

Therefore, (41) is the system of equations we must solve (for $\vec{x}$) in order to find our optimal point $\vec{x}^*$. But here's the thing - as we know, the gradient operator takes the derivative with respect to each x_i (i.e. $\nabla = [\partial_{x_1}, \partial_{x_2}, ...]$), but if you look at the functions h_m written out explicitly in (36)-(38), you see that each x_i appears only once throughout all of the h_m's. For example, for $i = 9$ we will have

$$\partial_{x_9} \left(\sum\nolimits_m \lambda_m \, h_m(\vec{x}) \right) = \lambda_2 \, \partial_{x_9} \Big(h_2(\vec{x}) \Big) \tag{42}$$

because x_9 doesn't appear in h_1 or h_3, rather it only appears in h_2. I'm going to tediously write out a similar expression for all i. What we find is that equation (41) is equivalent to the following system:

$$\left\{ \begin{aligned} w_1 + \mu_1 &= \lambda_1 \partial_{x_1} \Big(h_1(\vec{x}) \Big) \\ w_2 + \mu_2 &= \lambda_1 \partial_{x_2} \Big(h_1(\vec{x}) \Big) \\ w_4 + \mu_4 &= \lambda_1 \partial_{x_4} \Big(h_1(\vec{x}) \Big) \\ w_5 + \mu_5 &= \lambda_1 \partial_{x_5} \Big(h_1(\vec{x}) \Big) \\ w_8 + \mu_8 &= \lambda_2 \partial_{x_8} \Big(h_2(\vec{x}) \Big) \\ w_9 + \mu_9 &= \lambda_2 \partial_{x_9} \Big(h_2(\vec{x}) \Big) \\ w_{11} + \mu_{11} &= \lambda_2 \partial_{x_{11}} \Big(h_2(\vec{x}) \Big) \\ w_{12} + \mu_{12} &= \lambda_2 \partial_{x_{12}} \Big(h_2(\vec{x}) \Big) \\ w_{13} + \mu_{13} &= \lambda_3 \partial_{x_{13}} \Big(h_3(\vec{x}) \Big) \\ w_{15} + \mu_{15} &= \lambda_3 \partial_{x_{15}} \Big(h_3(\vec{x}) \Big) \\ w_{16} + \mu_{16} &= \lambda_3 \partial_{x_{16}} \Big(h_3(\vec{x}) \Big) \\ w_{18} + \mu_{18} &= \lambda_3 \partial_{x_{18}} \Big(h_3(\vec{x}) \Big) \end{aligned} \right\} \tag{43}$$

Sorry about that. But here comes the main problem - what I keep finding is that there is an inherent contradiction in this system, and thus the system cannot be solved.

To see why, this I need to define the notion of *conjugate pairs*. For any market m and any token k, let the phrase *conjugate pair* refer to the pair of variables $\big\{ \big(\vec{t}_m\big)_k, \big(\vec{r}_m\big)_k \big\}$. In other words, it's the tendered/received pair for the same token on the same market. In terms of the x_i variables, we see the conjugate pairs in our current example are (consult the table on page 4 for help):

$$\{x_1, x_4\}, \{x_2, x_5\}, \{x_8, x_{11}\}, \{x_9, x_{12}\}, \{x_{13}, x_{16}\}, \{x_{15}, x_{18}\} \tag{44}$$

Now, if we assume that there is a non-zero solution, then *at least one of these pairs must be completely non-zero.* Is this assumption correct? Doesn't it seem like a non-zero tendered amount **must** correspond with a non-zero received amount in the same pair?? But it turns out that if we assume there is a non-zero conjugate pair of variables in our final solution, then this leads to a contradiction.

Let me show you. Without loss of generality, suppose that our non-zero conjugate pair is $\{x_1, x_4\}$. Now consider the two equations from (43) corresponding to $i = 1$ and $i = 4$:

$$w_1 + \mu_1 = \lambda_1 \partial_{x_1} \Big(h_1(\vec{x}) \Big) \tag{45}$$

$$w_4 + \mu_4 = \lambda_1 \partial_{x_4} \Big(h_1(\vec{x}) \Big) \tag{46}$$

Let's explicitly compute the derivatives in (45) and (46). To do this, look back at the explicit expression for h_1 given in (36), which I'll remind you is the following:

$$h_1(\vec{x}) = \Big(R_{1A} + \gamma_1\, x_1 - x_4 \Big) \Big(R_{1B} + \gamma_1\, x_2 - x_5 \Big) - \Big(R_{1A} \Big) \Big(R_{1B} \Big) \tag{47}$$

Thus, we compute the derivatives:

$$\partial_{x_1} \Big(h_1(\vec{x}) \Big) = \big(\gamma_1 \big) \Big(R_{1B} + \gamma_1\, x_2 - x_5 \Big) \tag{48}$$

$$\partial_{x_4} \Big(h_1(\vec{x}) \Big) = \big(-1 \big) \Big(R_{1B} + \gamma_1\, x_2 - x_5 \Big) \tag{49}$$

We substitute these derivatives back into the equations (45) and (46):

$$w_1 + \mu_1 = \lambda_1 \big(\gamma_1 \big) \Big(R_{1B} + \gamma_1\, x_2 - x_5 \Big) \tag{50}$$

$$w_4 + \mu_4 = \lambda_1 \big(-1 \big) \Big(R_{1B} + \gamma_1\, x_2 - x_5 \Big) \tag{51}$$

Now, these two equations are meant to be solved together (technically they are part of the larger system (43), but we see that they separate from the larger system because they only reference each other). So, to solve them, let's divide (50) by (51):

$$\frac{w_1 + \mu_1}{w_4 + \mu_4} = -\,\gamma_1 \tag{52}$$

But! Remember our condition of **Complementary Slackness**, equation (30). Specifically, since we are assuming that $\{x_1, x_4\}$ are non-zero, then this means that $\{\mu_1, \mu_4\}$ *must be zero*! Thus, (52) reduces to the following:

$$\frac{w_1}{w_4} = -\,\gamma_1 \tag{53}$$

But go back to the definition of $\vec{w}$ in (34), and we see that $w_1 = u_A$ and $w_4 = -u_A$ (if you actually go check (34), remember that we eliminated the inactive indices already, so w_4 is in the third position). Plugging this into (53) gives us

$$1 = \gamma_1 \tag{54}$$

So, in order for a non-zero solution to exist, we must have $\gamma_1 = 1$, which translates to fees being 0%.

ARE YOU SHITTING ME???

Recall that γ_m isn't even supposed to be a variable, it's fixed input data. So the system of equations has absolutely no right to demand that γ_m be anything at all. This is a contradiction. If we started with $\gamma = 0.997$ (a fee of 0.3%), then the system of equations spits out $1 = 0.997$. Fat chance, I say. Something is fucked up.

I guess there are tons of places where I might have made a mistake.
Do you see anything wrong though???